1互联网

2019年3月10日 11:13

-
- 信息时代

1- 21世纪是以网络为核心的信息时代，特征有数字化、网络化、信息化

2- 三大网：电信网、有线电视网、计算机网

3- Internet：由数量极大的各种计算机网互连起来的网

1. 重要基本特点：connectivity连通性、resource sharing资源共享

-
- 互联网

1- 网络的网络

1. 网络由若干node结点和连接这些结点的link链路组成

(1) 结点：计算机、集线器、交换机、路由器等

2. interconnection：互连
3. interconnection network：互联网络
4. internetworking：网际互连
5. internet、internetwork、interconnection network：互联网，有时也可称互连网（不过全国科学技术名词审定委员会推荐音译为因特网）

6. 网络把许多计算机连接在一起，互联网通过路由器把许多网络连接在一起

(1) 与网络相连的计算机常称为host主机

2- 发展阶段

1. 最早的网络

(1) 1969年美国建立的分组交换网ARPANET被认为是Internet前身

(2) 1983年TCP/IP成为其标准协议

(3) 1990年ARPANET实验任务完成，正式关闭

(4) internet互连网是通用名词，泛指计算机网络互连成的计算机网络

(5) Internet互联网是专用名词，指全球最大的开放的互连网，采用TCP/IP协议

2. 三级结构

(1) 1985年美国National Scientific Foundation建立了NSFNET

(2) 分为主干网、地区网、校园网或企业网（覆盖大学和研究机构）

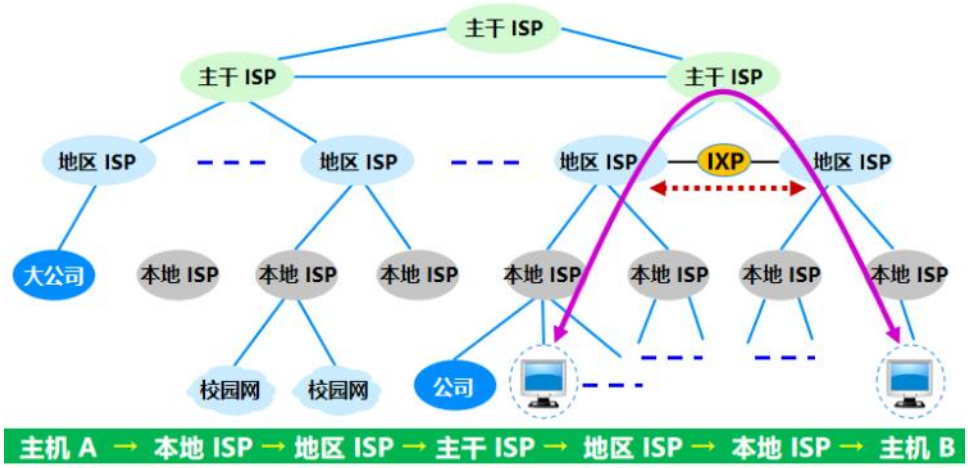
(3) 1993年美国政府不再负责互联网运营Internet

3. 多层次ISP结构

(1) Internet Service Provider互联网服务提供商ISP销售IP使用权

(2) Internet eXchange Point互联网交换点IXP允许两个网络直接相连并交换分组，常采用局域网互连的、数据链路层的网络交换机

(3)



(4) 欧洲原子核研究组织CERN开发的world wide web万维网www被广泛使用在互联网

3- 标准化工作

1. 1992年Internet Society互联网协会ISOC成立，管理互联网
2. ISOC的下属技术组织Internet Architecture Board互联网体系结构委员会IAB负责管理互联网协议（A曾经为Activities），它又下设两个工程部
   (1) Internet Engineering Task Force互联网工程部IETF负责短期开发
      1) 是由众多working group工作组gp组成的forum论坛
      2) 由Internet Engineering Steering Group互联网工程指导小组IRSG管理，研究若干area（主要是协议）的短、中期工程问题
   (2) Internet Research Task Force互联网研究部IRTF负责长期研究
      1) 是由Research Group研究组rp组成的forum
      2) 由Internet Research Steering Group互联网研究指导小组IRSG管理，研究协议、应用、体系结构等
   (3) 互联网标准以Request For Comments请求评论RFC的形式发表
      1) Internet Draft草案：6个月的有效期，不算RFC文档
      2) Proposed Standard建议标准：成为RFC文档，已有近万套
      3) Internet Standard标准：正式成为标准，已有近百套，可关联多个RFC文档
         -
         - 组成
   (1) 边缘部分：所有连接在互联网上的主机，用户可直接使用
   (2) 核心部分：大量网络和连接这些网络的路由器，向边缘提供服务

1- 边缘

(1) 主机又称为end system端系统

(2) 计算机之间通信：主机间的进程之间的通信

1. Client/Server客户/服务器方式C/S：client进程主动发起请求，server进程被动等待处理请求
2. peer-to-peer model对等连接方式P2P：双方可互相访问对方文件

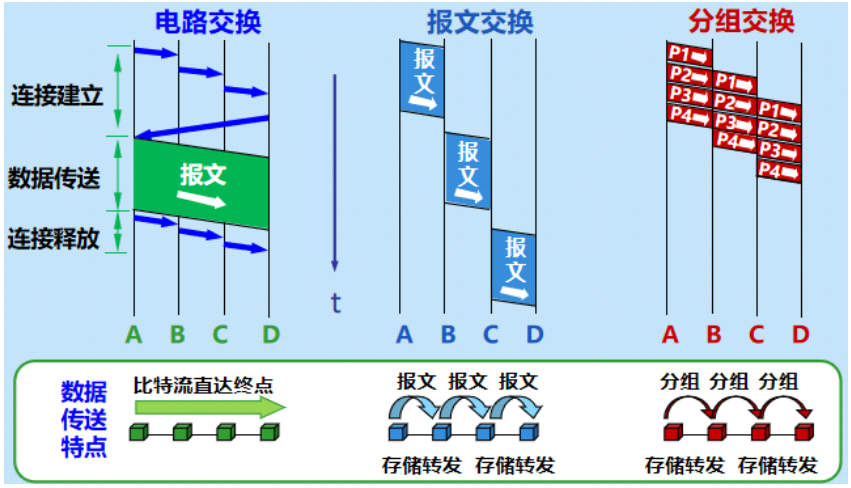
2- 核心

(1) router路由器是一种专用计算机

(2) router实现了packet switching分组交换，即转发收到的分组

1. circuit switching电路交换，专用物理通路法
   (1) 电路交换三步（电话机双绞线的交换）
      1) 建立连接、通话、释放连接
   (2) 通话期间，两个用户始终占用端到端的通信资源
   (3) 线路传输效率低
2. message报文交换存储转发技术
   (1) message报文：待发送数据
   (2) 加上控制信息，放在header首部，称为packet分组或包
   (3) 由报文交换中心操作员手动用相应发包机转发
   (4) 转发过程会产生时延，控制信息也会带来overhead开销
3. 和packet分组交换，存储转发技术
   (1) 主机负责处理信息，路由器负责自动转发分组
   (2) 路由器作为结点，构成了链路
   (3) 路由器通过protocol协议自动找最合适的链路
   (4) 分组交换优点

高效	动态分配传输带宽，对通信链路逐段占用，提高了channel信道利用率
灵活	为每个分组独立选择最合适的转发路由
迅速	以分组作为传送单位，可以不先建立连接就能向其他主机发送分组
可靠	保证可靠性的网络协议；分布式多路由的分组交换网有很好的生存性



-
- 在我国的发展

1- 1980年铁道部便开始建设广域网

2- 1989年11月我国第一个共用分组交换网CNAPC建成运行

3- 80年代起，公安、银行、军队等部门也开始建立专用广域网，许多单位也安装了大量局域网，价格便宜，所有权和使用权都属于本单位，易开发，易管理

4- 五大公用网

1. 中国电信互联网 CHINANET（也就是原来的中国公用计算机互联网）
2. 中国联通互联网 UNINET
3. 中国移动互联网 CMNET
4. 中国教育和科研计算机网 CERNET建于1994，我国首个IPv4主干网
5. 中国科学技术网 CSTNET

5- ChiNa Network Information Center中国互联网络信息中心 CNNIC每年两次公布我国互联网的发展情况

-
- 类别

1- 常见的定义：由一些通用的、可编程的硬件互连而成的，而这些硬件并非专门用来实现某一特定目的（例如，传送数据或视频信号）。这些可编程的硬件能够用来传送多种不同类型的数据，并能支持广泛的和日益增长的应用

2- 分类

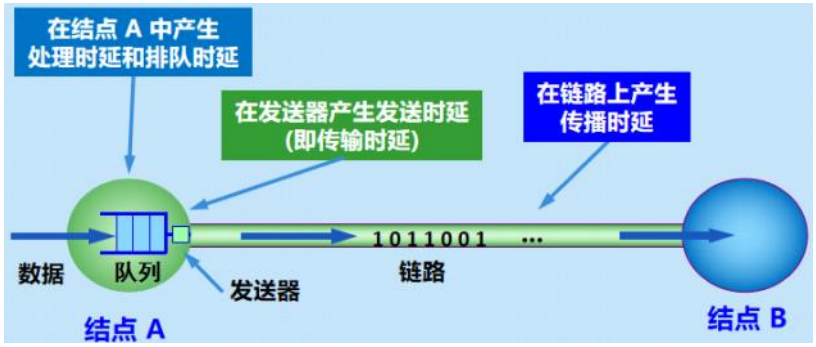
1. 按作用范围分类
   (1) Wide Area Network广域网WAN，或long haul远程网。几千公里
   (2) Metropolitan Area Network城域网MAN。几至几十千米
   (3) Local Area Network局域网LAN，常指校园网、企业网。一千米内
   (4) Personal Area Network个人局域网PAN。几米
   (5) 另外，CPU距离小于一米的一般不视作网络，只称为多处理机系统
2. 按使用者分类
   (1) public network公用网或公众网，缴费就能用
   (2) private network专用网，只给单位内部使用
3. 用来把用户接入到互联网的网络
   (1) Access Network接入网AN，又称为本地接入网或居民接入网
   (2) AN既不属于互联网的核心部分，也不属于互联网的边缘部分
   (3) 边缘路由器：某个用户端系统到互联网中的第一个路由器
   (4) 接入网是边缘路由器之间的一种网络，一般是局域网
   (5) 早期用户需要电话线拨号接入互联网，当时没有AN
   (6) 多种宽带接入技术的出现使宽带接入网成为新的热门课题
      1)
      2)
      3)
      4)
      5)
      6)
      7) ----------我是底线----------

1性能、体系

2019年3月10日 11:13

-
- 性能

1- 性能指标

1. 传输速率
   (1) bit比特是binary digit的简称，用作信息论中信息量的单位
   (2) data rate数据率或bit rate比特率为网络性能重要指标，单位b/s
   (3) 单位前缀：kilo千、mega兆、giga吉、tera太10^12、peta拍10^15、exa艾10^18、zetta泽10^21、yotta尧10^24
   (4) 往往不以平均速率而以额定速率或达标速率代表速率
2. bandwidth带宽或通频带，衡量能传输的最高数据率
   (1) 频域称谓：指某个信号具有的频带宽度，单位赫兹
   (2) 时域称谓：某通道中能通过的最高数据率，单位b/s
3. throughput吞吐量：单位时间内通过的实际数据率，单位b/s或帧/s
4. delay或latency时延或延迟：数据在端之间传输所需时间
   (1) transmission发送时延或传输时延：主机或路由器发送全部数据所需时间，即发第一个bit开始到最后一个bit结束的时间。高速链路一般是指发送时延少的链路
   (2) propagation传播时延：电磁波在信道中传播一定距离花的时间
   (3) 处理时延：分析首部、提取数据、差错检验、挑选路由
   (4) 排队时延：等待被路由器分析或转发的时间，取决于通信量
   (5) 

5. 时延带宽积：传播时延x带宽，表示链路可容纳的数据量，单位比特
6. Round-Trip Time往返时间RTT：发送数据开始到接收到确认信息的时间
   (1) 有效数据率=数据长度/发送时间+RTT
7. 利用率U
   (1) 信道利用率：信道被利用时间占比
   (2) 网络利用率：信道利用率的加权平均值
   (3) 时延D=D空闲/1-U
   (4) 一般利用率大于一半时需要考虑增大带宽

2- 非性能特征：费用、价格、标准化、可靠性、可扩展性和可升级性，易管理维护

-

◆ 体系结构

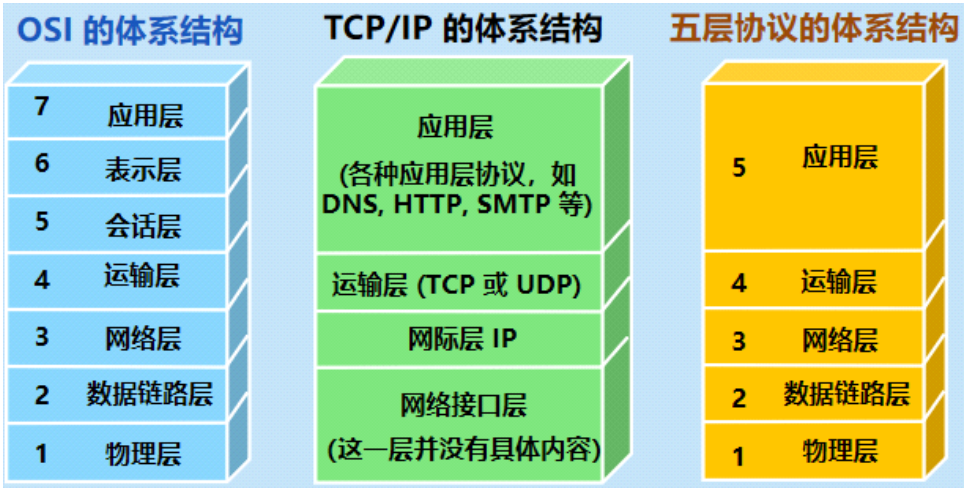
1- 体系结构的形成

1. 建立通路后的工作：active激活通路（发出信令确保能发送数据）、建立识别目标机的方法、确认目标机开机且连接到网络、确认做好接受准备、格式转换、差错处理等
2. 最早的体系结构是IBM在1974提出的System Network Architecture系统网络体系结构SNA，之后各种公司都相继推出不同的体系结构
3. 国际标准化组织ISO于1977推出了Open System Interconnection Reference Model开放系统互连基本参考模型OSI/RM
4. OSI是法律上理想的标准，但事实上的标准是TCP/IP
   (1) OSI 的专家们在完成 OSI 标准时没有商业驱动力
   (2) OSI 的协议实现起来过分复杂，且运行效率很低
   (3) OSI 标准的制定周期太长，按 OSI 标准生产的设备无法及时进入市场
   (4) OSI 的层次划分也不太合理，有些功能在多个层次中重复出现
5. OSI标准除了ISO外还有国际电报电话咨询委员会CCITT参与制订，CCITT后被International Telecommunication Union国际电信联盟ITU决定，与国际无线电咨询委员会CCIR合并成Telecommunication Standardization电信标准化部门TSS

2- 协议与划分层次

1. 解决数据交换同步问题的network protocol网络协议：
   (1) 语法：数据与控制信息的结构或格式
   (2) 语义：需要发出何种控制信息，完成何种动作以及做出何种响应
   (3) 同步：事件实现顺序的详细说明
2. 分层的优点：各层间独立、灵活、结构可分、易实现和维护、促进标准化
3. 各layer层需要完成的功能：差错控制、流量控制、分段和重装、复用和分用、建立和释放连接
4. architecture体系结构：各层及协议的集合，即网络各构件应完成的功能的精确定义，是抽象的
5. 计算机网络的实现是具体的，是计算机硬件和软件

3- 五层协议的体系结构（OSI和TCP/IP的折中版）

1. 

   (1) application应用层：进程间通信交互的规则，单位为message报文
      1) 如域名系统DNS、万维网HTTP、电子邮件SMTP
   (2) transport运输层：负责向进程通信提供通用的数据传输服务

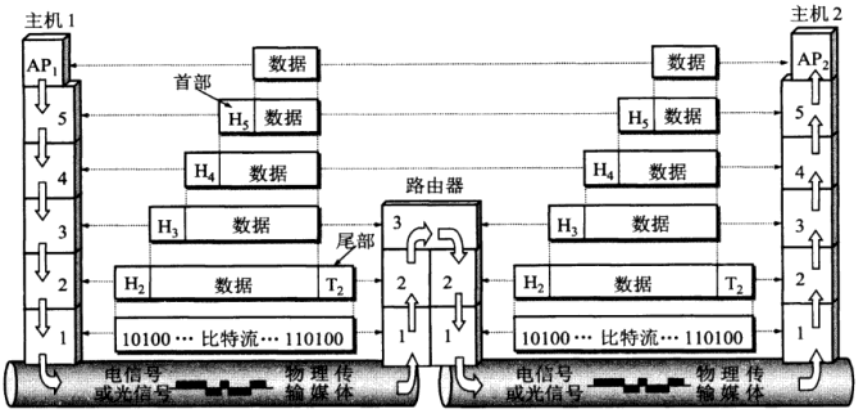
1) Transmission Control Protocol传输控制协议TCP：面向连接的可靠数据传输服务，单位是segment报文段
2) User Datagram Protocol用户数据报协议：无连接的best-effort尽力数据传输，单位是用户数据报
3) 注意TCP是transmission传输，而这层叫transport运输层

(3) network网络层：负责分组交换，单位是分组或包
1) 对应TCP/IP协议的网际层和IP数据包
2) UDP数据报和IP数据包不是一种单元，但分组可以泛指各种单元
3) 互联网是大量heterogeneous异构网络用router连接而成的
4) 互联网的网络层使用了Internet Protocol网际协议IP和多种路由选择协议，因而称为网际层或IP层

(4) data link数据链路层：负责将IP数据报组装成frame帧，单位是帧
1) 每帧都有数据和控制信息（同步信息、地址信息、差错控制等）

(5) physical物理层：负责制订物理媒体的使用标准，单位是比特
1) 但物理媒体本身不在物理层协议内

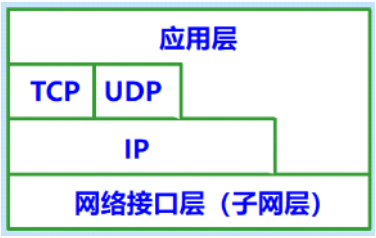
(6)



2. TCP/IP是指现在互联网使用的protocol suite协议族，TCP和IP是最重要的两个独立协议
3. peer layers对等层间传送数据的单位在OSI中被称为Protocol Data Unit协议数据单元PDU
4. 层次叠加的结构有时被称为protocol stack协议栈

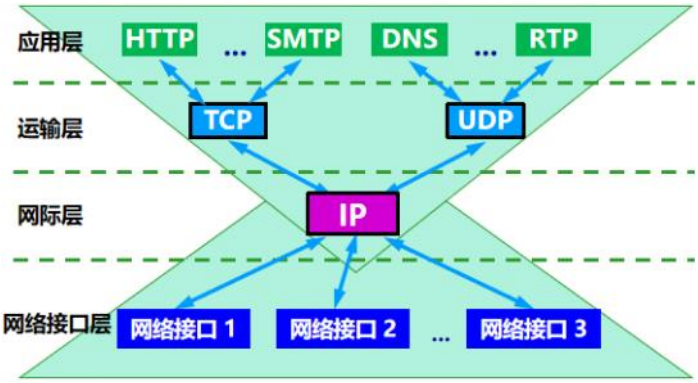
4- 实体、协议、服务和服务访问点
1. entity实体：泛指能发送或接收信息的硬件或软件进程
2. 协议是控制多个实体进行通信的规则的集合，使实体能向上层提供服务
3. 使用下层提供的服务需要交换一些命令，这些命令在OSI被称为服务原语
4. 相邻两层交换信息的地方被称为Service Access Point服务访问点SAP
5. OSI将上下层之间交换的数据单位称为Service Data Unit服务数据单元SDU
6. 有时多个SDU合为一个PDU，有时一个SDU可分为多个PDU
7. 协议需要考虑各种异常情况，不能假定一切都理想

5- TCP/IP的体系结构

1. 

   (1) 实际上的应用层演变成类似上图的分法

   (2) 此处的子网层指局域网和一些广域网（如ATM网），与子网划分无关

2. everything over IP：IP层为各种应用提供服务
3. IP over everything：IP协议可在各种网络构成的互联网上运行
4. 

   1)

   2)

   3)

   4)

   5)

   6)

   7) ----------我是底线----------

2物理层概念

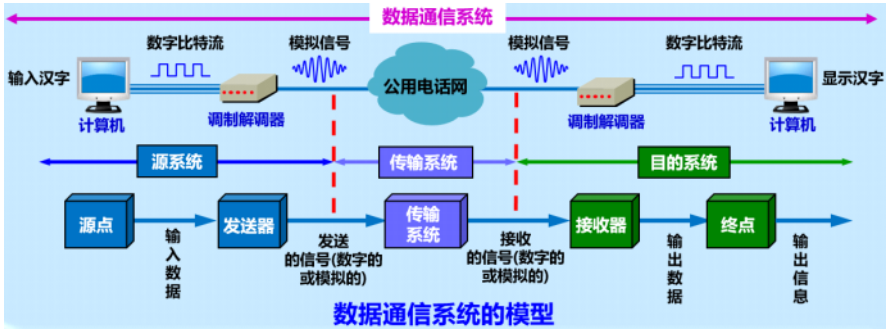
2019年4月7日　　0:18

-
- 基本概念

1. 物理层考虑如何连接各传输媒体上数据比特流，不考虑具体传输媒体
2. 物理层的作用是尽可能屏蔽不同传输媒体和通信手段的差异
3. 物理层协议/物理层规程procedure
4. 接口特性
   (1) 机械特性：接线器形状、尺寸、引脚数、排列等
   (2) 电气特性：电压范围
   (3) 功能特性：电压意义
   (4) 过程特性：不同功能的各种可能事件的出现顺序

-
- 通信基础

1- 通信系统：源系统、传输系统、目的系统

(1) 

数据通信系统的模型

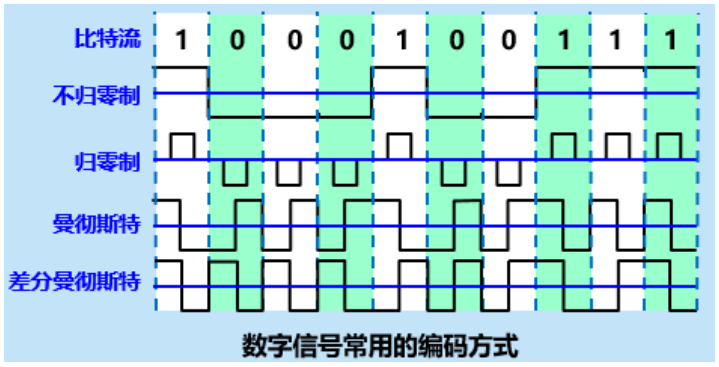
   (2) source信源/源点/源站：产生数据
   (3) 发送器/编码器：将源点的比特流编码后交给传输系统，如调制器
   (4) 接收器：将传输系统发来的信号转换回信息，如解调器
   (5) destination终点或目的地或信宿：将接收器的数据输出

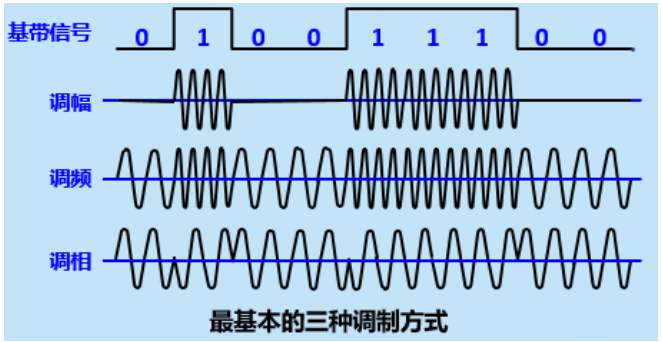
1. 常用术语
   (1) message消息：通信的目的，如语音、文字、图形、视频
   (2) data数据：运送消息的实体
   (3) signal信号：数据的电气或电磁的表现
   (4) 模拟信号/连续信号：参数取值连续的消息
   (5) 数字信号/离散信号：参数取值离散的消息
   (6) **码元**：代表不同离散数值的基本波形

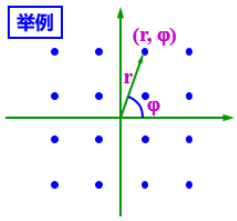
2- channel信道：一般指单向传输信息的媒体

   (1) 单向通信/单工通信：一方发一方接
   (2) 双向交替通信/半双工通信/广义单工通信：双方都能发，但不能同时发
   (3) 双向同时通信/全双工通信：可同时发送和接收，需要两条信道

1. 基带信号/基本频带信号：信源发出的信号，直接表达了信息，如计算机的10，电话的声音

2. 宽带信号：载波调制后的基带信号
   (1) 为了方便在信道传送，一般会调高频率范围，称这种转换为利用了carrier载波的**带通调制**，不搬移频段的称为基带调制
3. coding编码：将数据转化成数字信号
   (1) 数字发送器：将数字数据转化成数字信号
   (2) PCM编码器：将模拟数据转化成数字信号
      1) 抽样、量化、编码
   (3) 

数字信号常用的编码方式
      1) 自同步能力：从信号波形本身中提取信号时钟频率的能力
      2) Manchester编码及其差分形式和归零制都有自同步能力，信号频率也高于不归零制
4. modulation调制：将数据转化成模拟信号
   (1) 调制器：将数字数据转化成模拟信号
      1) 相位数*振幅数=信号数
   (2) 放大器调制器：将模拟数据转化成模拟信号
      1) 复用
   (3) 

最基本的三种调制方式
      1) 调幅(AM)：载波的振幅随基带数字信号而变化
      2) 调频(FM)：载波的频率随基带数字信号而变化
      3) 调相(PM)：载波的初始相位随基带数字信号而变化
   (4) Quadrature Amplitude Modulation正交振幅调制 QAM
      1) 

      2) 可供选择的相位有 12 种，而对于每一种相位有 1 或 2 种振幅可供选择。总共有 16 种组合，即 16 个码元
      3) 由于 4 bit 编码共有 16 种不同的组合，因此这 16 个点中的每个点可

对应于一种 4 bit 的编码。数据传输率可提高 4 倍

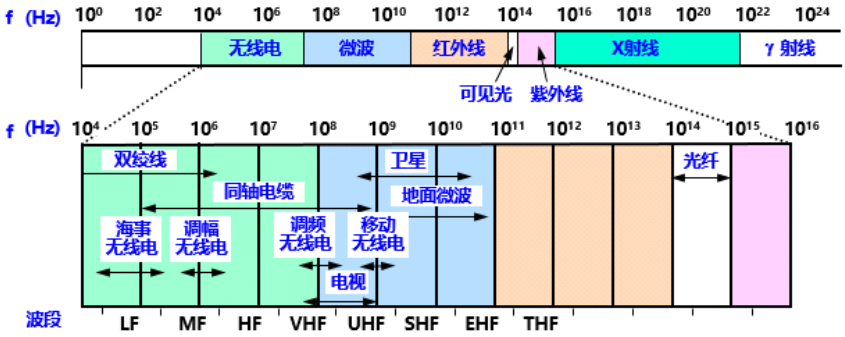
4) 但码元过多会导致识别困难和出错率增加

3- 信道的极限容量

1. **码元**：代表不同离散数值的基本波形
   (1) 对二进制信号来说，一码元=一比特
   (2) 对2^n进制信号来说，一码元=n比特
   (3) 码元传输速率/码元速率/波形速率/调制速率/符号速率
   (4) 单位Baud波特=码元/s
2. Nyquist奈氏准则：不考虑噪声，能避免码间串扰的码元传输上限
   (1) 码间串扰：接收端收到的信号波形中识别不出码元间的清晰界限
   (2) 码元传输的速率越高，或信号传输的距离越远，或传输媒体质量越差，在信道的输出端的波形的失真就越严重
   (3) 理想低通信道下的最高码元传输速率 = **2W** (Baud)
      1) 其中W为理想低通信道的带宽，单位Hz
3. Shannon**香农定理**：只考虑噪声的信道极限数据传输率
   (1) 信息传输速率/信息速率/比特率，单位比特/秒
   (2) 信噪比：信号平均功率与噪声平均功率之比，记为S/N，度量单位是分贝(dB)
      1) 实际计算时用的信噪比(dB) = 10 log10(S/N) (dB)
      2) 当S/N=10时，信噪比为10dB；S/N=1000时，信噪比为30dB
   (3) 信道的极限信息传输速率 C = **W log2(1+S/N)** (bit/s)
      1) 其中：W 为信道的带宽，单位 Hz；S 为信道内所传信号的平均功率；N 为信道内部的高斯噪声功率
   (4) 对应频带宽度已确定的信道，不能提高信噪比或码元传输速率时，就只能靠改变编码方法，使每个码元能携带更多比特的信息量

-
- 物理层下面的传输媒体/传输介质/传输媒介

电信领域使用的电磁波的频谱：

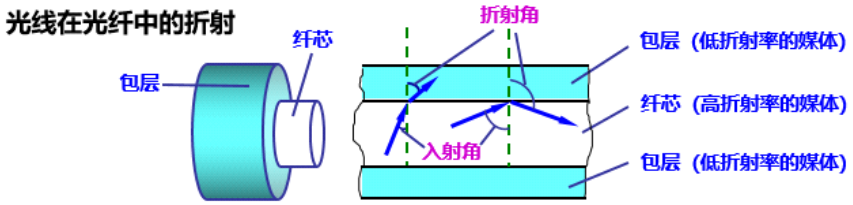


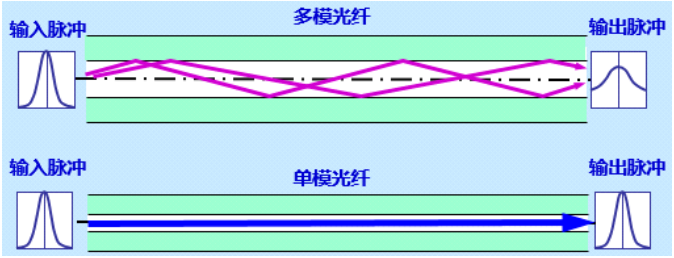
1. low、medium、high、very、ultra、super、extremely、tremendously

1- guided导引型传输媒体：电磁波沿固体媒介传播

1. 双绞线：两根互相绝缘的通道县并排twist在一起
   (1) 如电话用户线subscriber loop就是它
   (2) 模拟传输和数字传输都可以使用，通信距离一般为几到十几公里
   (3) Shielded Twisted Pair屏蔽双绞线 STP：带金属屏蔽层

(4) Unshielded Twisted Pair无屏蔽双绞线 UTP：较便宜

2. 同轴电缆：铜质内导体、绝缘层、网状外导体屏蔽层
   (1) 早期局域网用过，不过现在基本也用双绞线了
   (2) 现在主要用于有线电视网
3. 光缆：光纤通信的传输媒体，如上图，频率极高
   (1) 

   (2) 光线从高折射率的媒体射向低折射率的媒体时，其折射角将大于入射角。只要入射角足够大，就会出现全反射，光也就沿着光纤传输下去
   (3) 

   (4) 多模光纤：允许不同角度入射的光线同时传输
   (5) 单模光纤：直径小到约一个光的波长，光线几乎只能沿直线传
   (6) 优点
      1) 通信容量非常大
      2) 传输损耗小，中继距离长
      3) 抗雷电和电磁干扰性能好
      4) 无串音干扰，保密性好
      5) 体积小，重量轻

2- 非导引型传输媒体：自由空间传播无线电波
1. 短波通信/高频通信主要靠电离层的反射，通信质量较差，传输速率低
   (1) 多径效应：沿不同反射路径到达同一接收点，由于衰减和时延不同，使最后的合成信号失真很大
2. 微波在空间主要是直线传播，所以需要地面微波接力通信或卫星通信
   (1) 适用于电话、电报、图像、数据等
3. 优点
   (1) 频段宽，容量大
   (2) 该频段受到的干扰一般不大
   (3) 投资少见效快，能跨山河
4. 缺点
   (1) 相邻站必须能直视
   (2) 会受恶劣天气影响
   (3) 隐蔽性和保密性较差
   (4) 中继站维护耗费大
   (5) 卫星通信传播时延大

1)
2)
3)
4) ----------我是底线----------

2物理层技术

2019年5月3日 15:03

-
- 信道复用multiplexing

1. 复用器multiplexer和分用器demultiplexer之间的即为共享信道

1- 频分、时分、统计时分

1. Frequency Division Multiplexing频分复用FDM：用户同时占用不同带宽的资源，此处带宽指频率带宽
2. Time Division Multiplexing时分复用TDM：用户不同时占用同一带宽资源
   (1) 一般是周期性地给各用户使用，该周期可称为isochronous等时
   (2) 更利于数字信号传输
3. Statistic TDM统计时分复用STDM：动态分配时隙
   (1) 需要一个concentrator集中器，缓存不同用户的信息，再分成一个个帧一起发送出去，此处帧不是数据链路层的帧
   (2) 又称为异步时分复用，相对的可将前一种称为同步时分复用
   (3) 若各用户不是间歇性工作，就不能体现出和同步时分复用相比的优点

2- Wavelength Division Multiplexing波分复用WDM

1. 即光波的频分复用，用一根光纤传输多个光载波信号

3- Code Division Multiplexing码分复用CDM或Code Division Multiple Access码分多址CDMA：各用户选择不同的码型

1. 频谱类似白噪声，有很强的抗干扰能力，不易被敌人发现
2. 每一个比特时间划分为 m 个短的间隔，称为chip码片
   (1) 每个间隔对应1个1或-1
   (2) 因为 1 bit要转换成 m bit的码片，发送数据率必须也提高成m倍才能保持原有的信息率，这是一种spread spectrum扩频通信
   (3) Direct Sequence Spread Spectrum直接序列扩频DSSS
   (4) Frequency Hopping Spread Spectrum跳频扩频FHSS
   (5) 码片序列属于DSSS
3. 每站被指派一个唯一的m bit chip sequence码片序列
   (1) 想发送1时发送该序列
   (2) 想发送0时发送各元素乘以了-1后的该序列
4. 各站被分配到的码片序列必须各不相同且互相orthogonal正交
   (1) 正交即inner product内积=0
   (2) 利用正交性质和向量和的内积性质，收到任意个向量的叠加信号时，与每个站的码片序列求一个内积，为0时表示该站没有发送
   (3) 为m时（即规格化内积为1时）表示发送了1，为-m时表示发送了0

-
- 数字传输系统

1- 早期电话网

1. 从市话局到用户电话机的用户线是采用最廉价的双绞线电缆，而长途干线采用的是

频分复用 FDM 的模拟传输方式

2. 与模拟通信相比，数字通信无论是在传输质量上还是经济上都有明显的优势
3. 脉码调制 PCM 体制最初是为了在电话局之间的中继线上传送多路的电话，目前，长途干线大都采用时分复用 PCM 的数字传输方式

2- 由于历史原因，PCM有两种互不兼容的标准

1. 速率标准不统一
   (1) 北美日本T1 的速率是 1.544 Mbit/s
   (2) 欧洲中国E1 的速率是 2.048 Mbit/s
   (3) 如果不对高次群的数字传输速率进行标准化，国际范围的基于光纤高速数据传输就很难实现
2. 不是同步传输
   (1) 过去，为了节约经费，各国的数字网主要采用准同步方式
   (2) 当数据传输的速率很高时，收发双方的时钟同步就成为很大的问题

3- Synchronous Optical Network同步光纤网 SONET和SDH

(1) Synchronous Transport Signal第 1 级同步传送信号 STS-1，传输速率是 51.84 Mbit/s
(2) Optical Carrier第 1 级光载波 OC-1

1. ITU-T 以SONET 为基础，制订出国际标准Synchronous Digital Hierarchy同步数字系列 SDH
2. SONET和SDH的意义
   (1) 使不同的数字传输体制在 STM-1 等级上获得了统一
   (2) 第一次真正实现了数字传输体制上的世界性标准
   (3) 已成为公认的新一代理想的传输网体制
   (4) SDH 标准也适合于微波和卫星传输的技术体制

- ◆
- ◆ 有线宽带接入技术

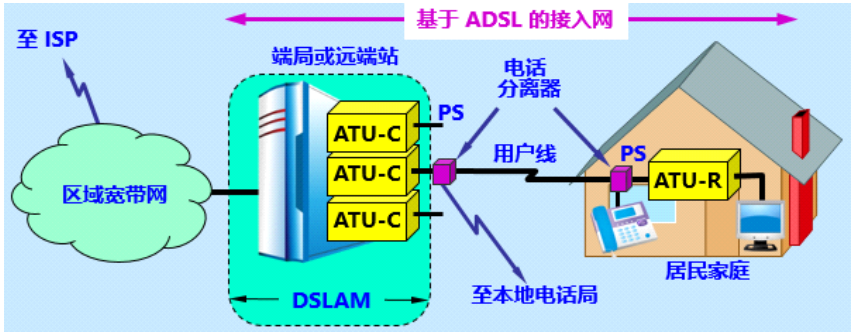
1. 美国联邦通信委员会FCC对宽带的最新定义：
   (1) 宽带下行速率要达到 25 Mbit/s
   (2) 宽带上行速率要达到 3 Mbit/s

1- Asymmetric Digital Subscriber Line非对称数字用户线 ADSL

1. 即：改造模拟电话用户线，使其能承载宽带业务
   (1) 把 0~4 kHz 低端频谱留给传统电话使用，把原来没有被利用的高端频谱留给用户上网使用
2. 传输距离取决于数据率和线径（用户线越细，信号传输衰减就越大）
3. 最高数据传输速率与实际的用户线上的信噪比密切相关
4. Discrete Multi-Tone离散多音调DMT：频分复用不同频段的信号
   (1) 一般用户是上行信道分配到25个子信道，下行分配到249个
5. 基于ADSL的接入网
   (1) DSL Access Multiplexer数字用户线接入复用器DSLAM
   (2) Access Termination Unit接入端接单元ATU，又称为ADSL调制解调器，是DSLAM的组成部件
   (3) ATU需要成对使用，Central端在端局使用，Remote端在用户处使用

(4) 用户电话通过电话Splitter分离器连接ATU-R，经线连接到ATU-C

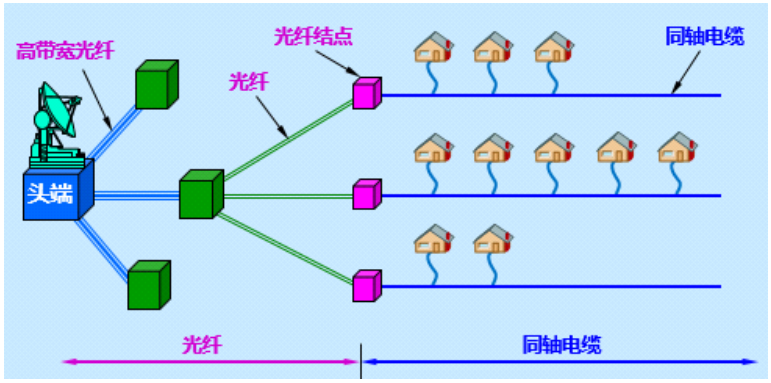
(5)



6. ADSL二代新标准的改进
   (1) 通过提高调制效率得到了更高的数据率：频谱范围从 1.1 MHz 扩展至 2.2 MHz，下行速率可达 16 Mbit/s（最大传输速率可达 25 bit/s），而上行速率可达 800 kbit/s
   (2) 采用了Seamless Rate Adaptation无缝速率自适应技术 SRA，可在运营中不中断通信和不产生误码的情况下，自适应地调整数据率
   (3) 改善了线路质量评测和故障定位功能，能提高网络的运行维护水平

2- Hybrid Fiber Coax光纤同轴混合网HFC网
1. 是在覆盖面很广的有线电视网 CATV 的基础上开发的一种居民宽带接入网
   (1) 除可传送 CATV 外，还提供电话、数据和其他宽带交互型业务
   (2) 是树形拓扑结构的同轴电缆网络，它采用模拟技术的频分复用对电视节目进行单向传输
   (3) 将原 CATV 网中的同轴电缆主干部分改换为光纤，使用模拟光纤技术
   (4) 在模拟光纤中采用光的振幅调制AM，比使用数字光纤更为经济

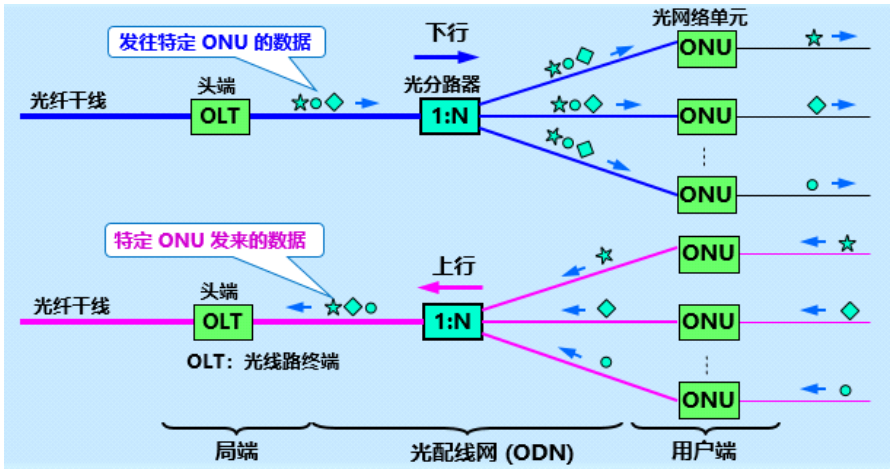
(5)



2. 模拟光纤从头端连接到fiber node光纤结点，即Optical Distribution Node光分配结点
   (1) 在光纤结点中，光信号被转换为电信号，之后就是同轴电缆了
   (2) 光纤结点到头端典型距离25km，到用户距离则不超过3km
   (3) 不需要成对使用，只需在用户端安装cable modem电缆调制解调器
3. User Interface Box用户接口盒UIB提供三种连接
   (1) 同轴电缆连接到set-top box机顶盒，在连接到电视机
   (2) 双绞线连接到用户电话机
   (3) 电缆调制解调器连接到用户计算机
4. 美国有线电视实验室制定了DOCSIS规约
   (1) 一般下行可达到数Mb每秒，最高小数十Mb每秒

(2) 大量用户同时上网时网速会很慢

3- Fiber To The X 光纤到X技术 FTTx

1. 可以到Home户、Building楼、Curb路边、Zone小区、Floor层、Office办公室、Desk桌面
   (1) 其实现在陆地上长距离运输媒体基本都换成光缆了，只是在用户端接口可能还是在用铜缆
   (2) Optical Distribution Network光配线网ODN负责把光纤的数据共享给各家庭用户
2. Passive Optical Network无源光网络PON
   (1) 因无需电源，基本不需维护，而广为使用
   (2) Optical Line Terminal光线路终端OLT负责连接光纤干线，再发给：
   (3) 光Splitter分路器负责广播给各用户端的：
   (4) Optical Network Unit光网络单元ONU，将与自己特有标识相同的光信号接收下来，转换成电信号，ONU所在位置即FTTx的x
   (5) 

3. PON的两大种类
   (1) Ethernet PON以太网无源光网络EPON
      1) 指在链路层是用来以太网协议，用PON拓扑结构实现接入的网
      2) PON与以太网兼容性好，成本低，扩展性强，管理方便
   (2) Gigabit PON吉比特无源光网络GPON
      1) 采用Generic Encapsulation Method通用封装法GEM
      2) 可承载多业务，对各类型都能提供服务质量保证，是很有潜力的宽带光纤接入技术
      3)
      4)
      5)
      6)
      7) ----------我是底线----------

3数据链路层和点对点

2019年5月3日　　19:50

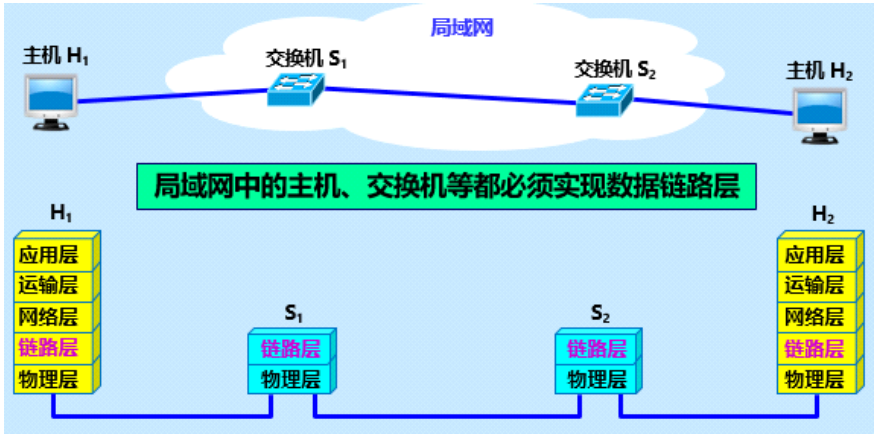
-
- 数据链路层

1- 数据链路和帧

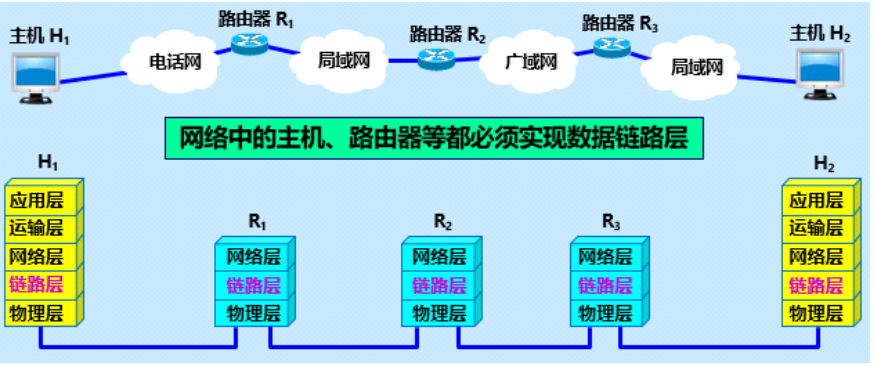
1. link链路：从结点到相邻结点的物理线路（也可以是无线的）
   (1) 是路径的组成部分，又称物理链路，对应将数据链路称为逻辑链路
2. data link数据链路=物理线路+通信协议（的硬件实现和软件实现）
   (1) **通信协议的硬软件实现**可统称为**网络适配器**
   (2) 这些通信协议也称procedure规程
3. frame帧：数据链路层基本协议数据单元
   (1) 而网络层的PDU是IP数据报（或称数据报、分组、包）
   (2) 帧中部分数据是需要上交给网络层的IP数据报

2- 数据链路层

1. 数据链路层通信步骤：

(1) 

   (2) 将网络层交下来的IP数据报封装成帧，添加首部和尾部
   (3) 将帧发送给目的结点的数据链路层
   (4) 目的结点确认帧无差错则取出IP数据报，上交给网络层，否则丢弃帧

(5) 

2. 两种信道
   (1) 点对点信道：一对一的点对点通信
   (2) 广播信道：一对多的广播信道

3- 三大基本问题

1. framing封装成帧：给数据报前后添加首部和尾部

(1) 首部和尾部的作用之一是帧定界，并初步判断帧内容是否完整

(2) 可打印字符：键盘47键内容及其shift后的内容+空格=95种字符

(3) 当数据内容都是可打印字符时，帧定界符可以简单地用Start Of Header和End Of Transimission联众控制字符表示（分别是二进制码0001和0100）

2. transparent transmission透明传输

(1) 指任何数据都能无差错地，感受不到数据链路层一般地传输

(2) 若数据非纯ASCII码文件，且恰出现了SOH或EOT，会判断失误

(3) 为防止以上这种错误影响透明传输，可以使用byte stuffing字节填充或称character stuffing字符填充，即在类似控制字符的数据前添加转义字符ESC（十六进制1B），并在ESC前也添加ESC（接收端再删除）

3. error detection差错检测

(1) 比特错误：1变成0或0变成1，与噪声有关

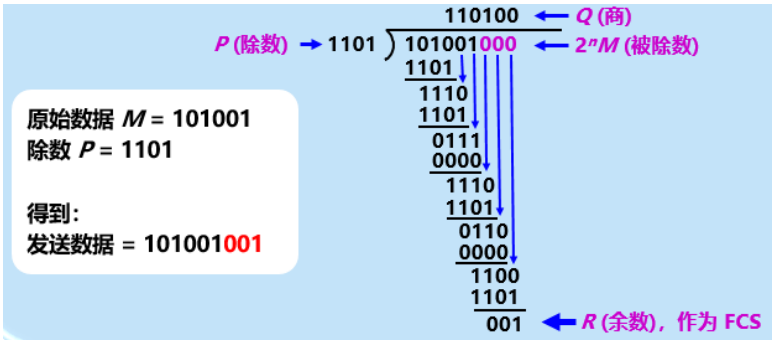
1) Bit Error Rate误码率BER：错误比特数占总比特数的比率

2) Cyclic Redundancy Check循环冗余检验法CRC：在数据后添加检测用的冗余码，该码又称Frame Check Sequence帧检验序列FCS

3) FCS求法：给M位原码乘以2^n（即左移n位），再除以一个实现预定号的除数P（生成多项式），得到的n位余数R即为FCS（此处除法中用到的减法是模2不借位减法，即异或）

4) CRC检验法：将带R的码除以P，得到余数R为0即默认无误

5)



(2) 非比特错误的传输错误：帧丢失、帧重复、帧失序

1) 解决方法：帧编号、确认、重传机制

2) 对于通信质量良好的优先传输链路，现在一般不要求确认和重传，改正差错的任务交给上层协议完成（如运输层的TCP）

☑3) 以上做法可以提高通信效率，但并不能视作可靠传输，真正的可靠传输协议在第五章

-
- 点对点协议Point-to-Point Protocol

1- PPP协议的特点

(1) 在通信线路质量较差的年代，普遍希望数据链路层能提供足够可靠的传输协议，当时使用的是High-level Data Link Control高级数据链路控制HDLC协议

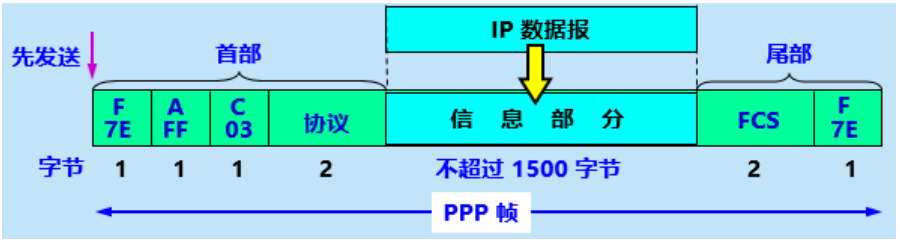
(2) 1994年起简单的PPP协议成为了互联网的正式标准

(3) PPP不支持多点线路通信，不支持单工通信和半双工通信，即只支持全双工链路的点对点通信

(4) 另外，因为TCP/IP族的可靠传输交给了TCP协议负责，所以PPP不需要考虑纠错、设置序号、流量控制

1. PPP协议应满足的需求
   (1) 首要需求：简单、互操作性高
   (2) 封装成帧：对帧定界符达成共识
   (3) 透明性：实现透明传输
   (4) 兼容网络层：在同一物理链路上支持多种网络层协议
   (5) 兼容物理链路：在串行或并行、同步或异步、高或低速、电或光、动态或静态（交换或非交换）的各种链路上都能运行（如1999年公布的PPP over Ethernet，使任何用户都能轻松连接上到ISP的宽带链路）
   (6) 差错检测：对有明显差错的帧立即丢弃
   (7) 检测连接状态：几分钟内自动检测一次工作状态是否正常
   (8) 最大传送单元MTU：设置Maximum Transmission/Receive Unit值
   (9) 网络层地址协商：使两个网络层实体能得知彼此的网络层地址，这对拨号连接的链路特别重要
   (10) 数据压缩协商：协商压缩算法，不过并不需要标准化压缩算法
2. PPP协议的组成
   (1) 将IP数据报作为信息封装到串行链路的方法，又分为支持无奇偶检验的8比特异步链路和面向比特的同步链路
   (2) 建立、配置和测试数据链路连接的Link Control Protocol链路控制协议LCP，包含一些选项供通信双方协商
   (3) 一套Network Control Protocol网络控制协议NCP，用于支持各种网络层协议

2- PPP协议的帧格式

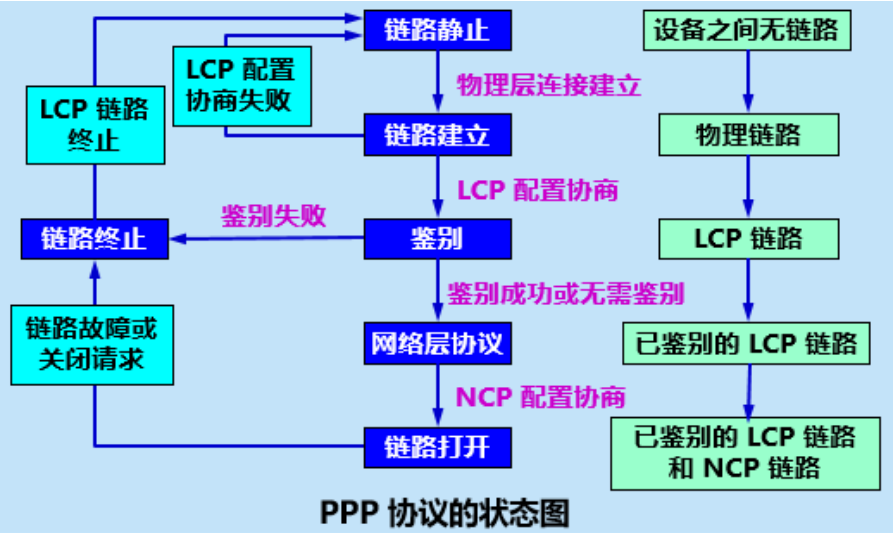
1. 各字段的意义
   (1)


   (2) F是帧定界符；AC是默认有的保留字段，暂时无义；FCS是检验码
   (3) 协议字段视信息部分而变：0x0021代表 IP 数据报；0x8021代表网络控制数据； 0xC021代表 PPP 链路控制数据；0xC023代表鉴别数据
2. 字节填充
   (1) 将转义字符定义为0x7D=01111101
   (2) 将信息部分的0x7E转换为0x7D5E
   (3) 将信息部分的0x7D转换成0x7D5D
   (4) 将信息字段的ASCII码转换为0x7D+一对应的其他字节
3. 零比特填充
   (1) 在SONET/SDH链路中，PPP协议会使用不逐个字符传输的并行传输
   (2) 此时需要避免出现连续6个1的，被误以为是边界符的情况

(3) 解决方法是每发现连续5个1就立刻发1个0，而接收方若未接受到连续六个1，就自动删除第5个1后面的0

3- PPP协议的工作状态

(1)



1. 工作起始和终止状态都是Link Dead链路静止状态，此时无物理层连接
2. 当电脑通过调制解调器呼叫路由器，并让路由器检测到调制解调器发出的载波信号后，PPP就开始进入Link Establish状态准备建立链路层LCP连接
   (1) 首先是通过Configure-Request配置请求帧协商配置选项
      1) Configure-Ack：确认接受所有选项
      2) Configure-Nak：确认理解选项但不能接受
      3) Configure-Reject：无法识别或不能接受选项，需要协商
   (2) 选项包括：最大帧长、authentication protocol鉴别协议、是否使用PPP帧的固定无义字段
3. 协商结束后进入Authenticate鉴别状态
   (1) 只允许传送LCP协议的分组、鉴别协议的分组及检测链路质量的分组
   (2) 如果使用Password Authentication Protocol口令鉴别协议PAP，则需要发起通信的一方发送身份标识和密码
   (3) 如果使用更复杂的Challenge-Handshake Authentication Protocol握手鉴别协议CHAP，则安全性会更高（协议原理是单向提出问题，等待对方返回正确答案后才允许继续连接，这个答案通常需要靠事先约定好的哈希函数来实现，这个三次握手过程在连接成功后也可以定期反复进行）
4. 鉴别失败后进入Link Terminate链路终止状态，回到静止态
5. 鉴别成功后进入Network-Layer Protocol网络层协议状态
   (1) 两端的NCP根据不同的网络层协议交换特定的网络控制分组
   (2) 如某一段使用IP协议时，两端都要配置好IP Control ProtocolIP控制协议IPCP，并定好帧内协议字段内容
   (3) 必要时还可协商压缩算法和IP首部要不要跳过无义字符
   (4) 可见PPP协议作为数据链路层协议的同时也考虑到了物理层和网络层
6. 协议配置完闭后链路进入Link Open链路打开状态
   (1) 两个端点可以开始彼此发送分组
   (2) 必要时也可发送Echo-Request回送请求分组和Echo-Reply会送回答分组，

用于检测链路的状态

7. 分组传输完后，可由任一端发送Terminate-Request终止请求，再确认Terminate-Ack终止确认后转入链路终止态
   (1) 链路故障后也会回到链路终止状态
   (2) 当调制解调器的载波停止后，链路会从终止态回到静止态
      1)
      2)
      3)
      4)
      5) ----------我是底线----------

3数据链路层广播信道

2019年5月3日 22:57

-
- 使用广播信道的数据链路层

(1) 局域网特点：为一个单位所拥有；地理范围和站点数目均有限

(2) 局域网优点

1) 有广播功能，从一个站点可很方便地访问全网

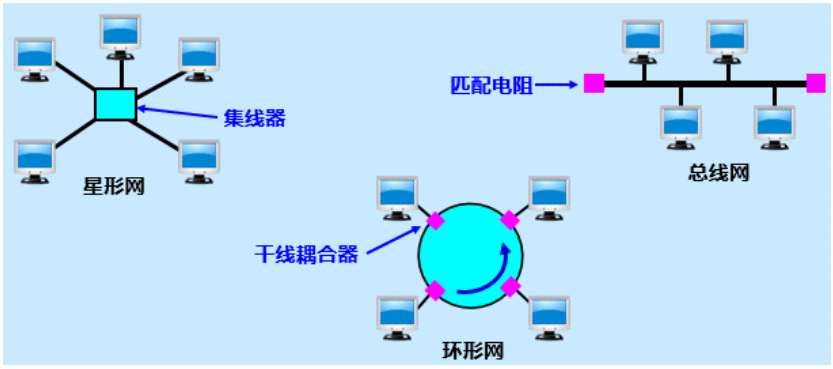
2) 局域网上的主机可共享连接在局域网上的各种硬件和软件资源

3) 便于系统的扩展和逐渐演变，各设备的位置可灵活调整和改变

4) 较高reliability可靠性、availability可用性和survivability生存性

(3) 局域网常见拓扑结构

1)



(4) 局域网传输媒体：一般是双绞线，仅当需要数Gb/s时才考虑光纤

(5) 局域网要考虑的问题其实不局限于数据链路层，物理层也有很多问题

(6) 如信道共享问题

1) 静态划分：即物理层介绍的方法，代价较高，并不适合局域网

2) 动态媒体接入控制，又称multiple access多点接入：信道并非在用户通信时固定分配给用户

a) 随机接入：任意发消息，发生碰撞冲突后同时失效，再由一定协议解决碰撞冲突，是本节主要内容

b) 受控接入：按一定规则地发消息，如分散控制的令牌环局域网和几种控制的多点线路polling探询或轮询

1- 局域网的数据链路层

(1) 以太网是美国Xerox施乐公司Palo Alto研究中心简称PARC研制的基带总线局域网，以历史上表示传播电磁波的以太命名

1. 以太网的两个标准

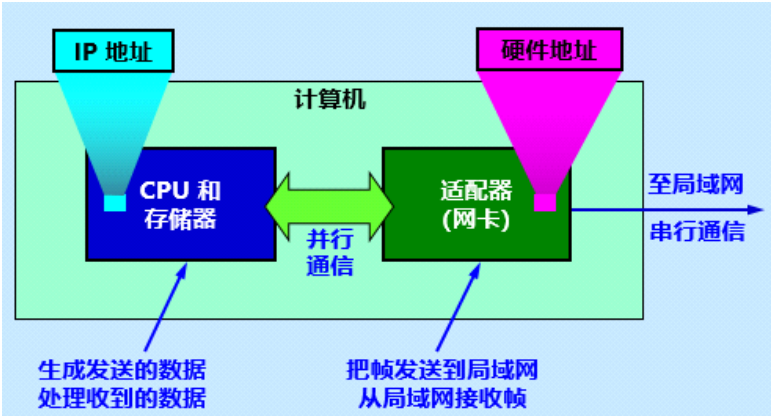
(1) DIX Ethernet V2 是世界上第一个局域网产品（以太网）的规约，是DEC、Intel、Xerox共同提出的

(2) 之后IEEE的802.3工作组也提出了对帧格式做出微小变动的新标准

(3) 两种标准基本兼容，都可视作广义上的“以太网”

(4) 为了适应商业竞争时的各种标准，802工作组将标准分为了Logical Link Control逻辑链路控制LLC和Medium Access Control媒体接入控制MAC两个子层，MAC对LLC是透明的

(5) 90年代后局域网市场基本被以太网垄断，甚至说等价于局域网，现在LLC协议已经甚至已不被适配器安装，本节也不再讨论

2. adapter适配器的作用：让计算机能与外界局域网连接
   (1) 最早的适配器是Network Interface Card网络接口卡/网卡NIC
   (2) 适配器和局域网的通信一般是通过电缆或双绞线的串行传输实现的
   (3) 适配器和计算机的通信一般是通过计算机主板的IO总线并行传输实现
   (4) 因此适配器还有个功能是**串并转换**和数据缓存
3. 适配器的工作
   (1) 一般计算机需要安装适配器的驱动程序，来控制适配器与内存的什么位置进行交互，而适配器需要实现以太网协议
   (2) 适配器的工作也不局限于是数据链路层，不过物理层的工作一般也被集成在了一起，不能分开考虑
   (3) 适配器接收和发送各种帧时一般CPU不参与干涉，适配器只在收到正确帧时会发出中断信号通知计算机，并交付给网络层；而计算机想要发出数据时，会通知适配器代为组装成帧并发送到局域网
   (4) 计算机的“硬件地址”存在适配器的ROM中，而软件地址则在计算机的存储器中
   (5) 

2- Ethernet 以太网 和 Carrier Sense Multiple Access with Collision Detection 载波监听多点接入/碰撞检测协议 CSMA/CD
   (1) 局域网上的计算机可称为主机、工作站、站点、站
   (2) 最初的以太网：所有计算机都连接在一根总线，因为当时都认为有源器件不可靠，无缘电缆最可靠
   (3) 当时的以太网，每当有总线上的计算机想发送数据，总线上其他所有主机都能检测到该数据，实现了无差别广播通信，如果需要一对一通信，可以通过在帧首部指定目的地址，只有硬件地址与目的地址相同的适配器可以接受该数据帧
1. 以太网的特点
   (1) 灵活的无连接工作方式：不对帧编号，不要求发回确认
      1) 即以太网只需提供“尽最大努力交付”=不可靠交付的服务，这是因为局域网信道质量值得信赖一般不考虑出错的问题，仅当CRC查出错时会抛弃帧，是否重传也交给高层决定
      2) 如TCP协议会要求重传，而以太网并不知道他需要进行重传，只负责传

TCP协议传下来的数据就完事了

3) 总线资源只能允许同时一台计算机发送数据，协调冲突的协议即为CSMA/CD，也就是说CSMA/CD一般工作在半双工通信时

(2) 发送数据信号默认为Manchester编码

1) 即根据一个码元的中间段是上升还是下降来判断10

2) 优点是遇到长串1或0也能轻松自同步，缺点是频带宽度翻倍了（即每秒传输的码元数量翻倍了）

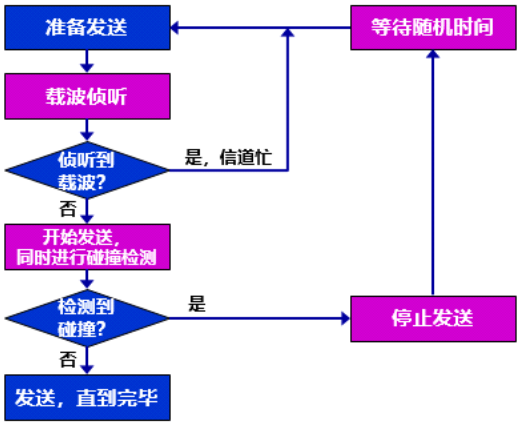
2. CSMA/CD的要点

(1) 多点接入MA：说明这是总线型网络

(2) 载波监听CS：其实并没有用载波，只是每个站在发送前后都必须时刻用电子技术检测信道是否空闲，因为只有不发生碰撞才能发送数据

(3) 碰撞检测CD：即一边发送数据一边要继续监听总线信号电压，超过一定门限值后说明有两个站在同时发送数据，并发生碰撞了

(4)



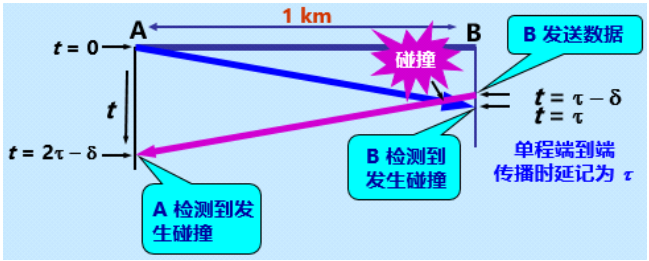
3. 关于碰撞检测

(1) 碰撞/冲突发生后会导致信号严重失真，无法识别，因此一旦检测出碰撞，就必须立即停止发送，等一段随机时间后再发送

(2) 电磁波在1km电缆的传播时延约5μs，每个站在发送数据后的几μs内并不确定有没有发生碰撞，这一特点称为以太网发送的不确定性

(3) 设单程端到端传播时延为t，则最坏情况在往返2t时才能检测到碰撞，称这个2t为content period争用期或collision window碰撞窗口

1)



(4) 经过争用期后就可确保这一帧没有发生碰撞了，相反的，如果收到帧长度<争用期*数据发送速率的帧，一般直接默认为是在争用期内由于冲突导致异常中止的无效帧，会被直接丢弃

(5) 发生碰撞的话就需要推迟或者说退避一段时间后，再重新发送，这个退避时间的算法如下：

4. truncated binary exponential backoff截断二进制指数避退算法

(1) 规定争用期为51.2μs，对于默认网速10Mb/s的以太网，这段时间内可以发

送512比特，有时可以直接以比特数衡量争用期长短

(2) 从整数集[0,1,…,2^k -1]中随机取出一个数r，则r倍争用期即为重传推后时间，其中k=min{ 重传次数,10 }

(3) 若k=16还没成功发出，则直接丢弃该帧，并向高层报告

(4) 这种随重传次数而增大推迟平均时间的退避称为动态退避

5. 争用期远长于实际端到端往返时间的原因

(1) 强化碰撞：检测到碰撞后，除了立即停止发送外，还要再发送32或48bit的jamming signal干扰信号，以便让其他用户明确发生了碰撞

(2) 因此实际上总线被占用的时间会达到3倍端到端传播时延

(3) 另外以太网还规定了帧间最小间隔9.6μs=96比特时间，以便清理缓存

6. CSMA/CD协议的工作要点

(1) 准备发送：适配器收到高层的分组，加上以太网的首部和尾部，组成以太网帧，放入缓存，开始检测信道

(2) 检测信道：信道忙时不停检测，连续96比特时间检测到空闲后才发送

(3) 发送后仍需不断检测信道，争用期内监听不到碰撞说明发送成功，回到1；检测到碰撞时应立即停止发送原数据，并发送一段干扰信号，等待一段按TBEB算法算出的时间后回到步骤2

(4) 注：发送后的争用期是需要保留帧数据的，争用期后确认无碰撞，即不需要重传后才可删除

(5) **最短帧长=传播速率/2倍传播时延**

3- 使用hub集线器的星形拓扑

1. 星形以太网

(1) 1990年IEEE推出了星形以太网标准10Base-T，10代表数据率为10Mb/s，BASE代表连接线上的信号为基带信号，T代表双绞线

(2) 现在粗缆和细缆的以太网都从市场上消失了

(3) 星形以太网中站点到集线器的距离一般不超过100m，这种以太网的出现是局域网史上极重要的一个里程碑，也确立了以太网的统治地位

(4) 另外，如果需要远程连接，也可以使用光纤做传输媒体，对应10BASE-F标准

2. 集线器的特点

(1) **物理上是星形网，逻辑上仍是总线网**，需要使用CSMA/CD

(2) 集线器有很多硬件端口或称为接口，很像是多接口的转发器，这些接口通过RJ-45插头连接计算机适配器（比电话机的RJ-11接口稍大）

(3) **集线器工作在物理层**，每个接口仅负责转发比特，不需要碰撞检测

(4) 集线器的专门芯片可抵消自适应串音回波，即每个比特在再生整形成强信号转发出去时不会干扰正在被接收的弱信号

3. stackable堆叠式集线器：将多个集线器堆叠起来使用

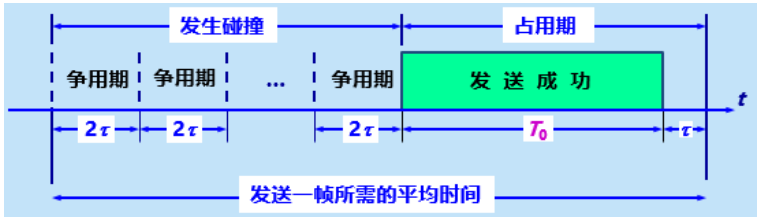
(1) 集线器一般有少量容错能力和网络管理功能，如一个故障适配器不断发送以太网帧时，集线器会在内部断开与其的连线

(2) 模块化机箱式智能集线器有很高的可靠性，网络功能以模块方式实现

(3) 各模块均可热插拔，不断电也可更换或增加新模块

(4) 集线器上的指示灯还可显示网络上的故障情况，方便网络管理

4- 以太网的信道利用率

1. 发生碰撞后，信道资源不能被任何站使用，也就是被浪费了
   (1) 

   (2) 帧的发送时间T0 = L/C s，其中帧长为 L bit，数据发送速率为 C bit/s
   (3) 设t是以太网单程端到端传播时延，即争用期的一半
   (4) 则成功发送一个帧需要占用信道的时间是 T0 + t，其中t是为避免碰撞而产生的被浪费的时间
2. 令参数a = t / T0，即单程端到端时延与帧发送时间T0之比
   (1) 易知a→ 0表示一发生碰撞就立即可以检测出来，并立即停止发送，信道利用率很高
   (2) a 越大，表明争用期所占的比例增大，每发生一次碰撞就浪费许多信道资源，使得信道利用率明显降低
   (3) 因此，为提高利用率，以太网的参数 a 的值应当尽可能小些
3. 改进思路
   (1) 数据率一定时，限制太网的连线的长度，使t减小
   (2) t一定时，增长以太网帧长，使T0增大
4. 信道利用率最大值（理想值）
   (1) $$S_{\max} = \frac{T_0}{T_0 + \tau} = \frac{1}{1 + a}$$
   (2) 据统计，当以太网的利用率达到 30% 时就已经处于重载的情况。很多的网络容量被网上的碰撞消耗掉了

5- 以太网的MAC层
1. MAC层的hardware硬件地址/物理地址/MAC地址
   (1) identification system标识系统中地址是作为识别系统的标识符（标准汉语词典推荐将标识zhi改成标志，为区分flag，此处才译作标识）
   (2) SHOC78的定义：名字指出所寻资源，地址指出资源在何处，路由告诉我们如何到达该处。按这种说法，局域网地址实为“名字”
      1) 原地更换新适配器也会更改MAC地址
      2) 不换适配器地移动计算机不会更改MAC地址
   (3) MAC地址是48位的二进制串
   (4) IEEE标准是允许16位MAC地址的，但为了方便标识，一般都用48位
2. IEEE的Registration Authority注册管理机构RA是局域网全球地址法定管理机构，负责分配地址字段的高24位
   (1) 生产局域网适配器的厂家需要购买前三个字节，作为Organizationally Unique Identifier组织唯一标识符OUI或称company_id公司标识符
   (2) 不过实际上一个公司可以买很多OUI，几个公司可以买同一个OUI
   (3) 后24位则称为Extended Unique Identifier扩展标识符
   (4) 拼接以后即得到了Extended Unique Identifier扩展唯一标识符EUI

(5) IEEE 规定地址字段的第一字节的最低位为 I/G 位，表示 Individual / Group，当 I/G 位 = 0 时，地址字段表示一个单站地址；当 I/G 位 = 1 时，表示组地址，用来进行多播

(6) IEEE 把地址字段第一字节的最低第 2 位规定为 G/L 位，表示 Global / Local，当 G/L 位 = 0 时，是全球管理（保证在全球没有相同的地址），厂商向 IEEE 购买的 OUI 都属于全球管理；当 G/L 位 = 1 时， 是本地管理，这时用户可任意分配网络上的地址

(7) 所有 48 位都为 1 时，为广播地址。只能作为目的地址使用

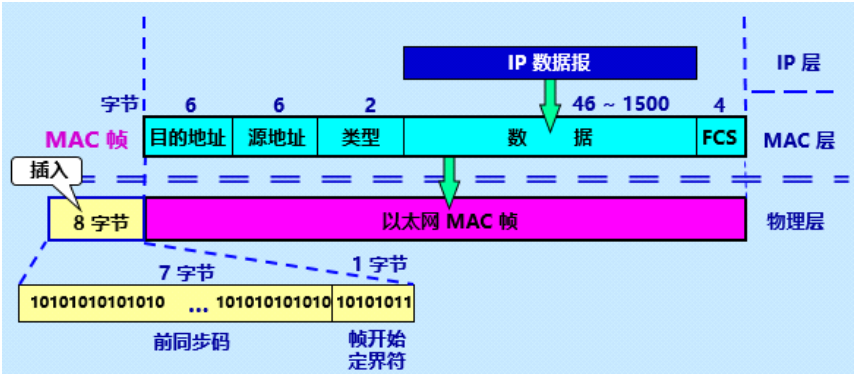
3. 适配器检查MAC地址：适配器从网络上每收到一个 MAC 帧就首先用硬件检查 MAC 帧中的 MAC 地址，如果是发往本站的帧则收下，然后再进行其他的处理。否则就将此帧丢弃

(1) “发往本站的帧”包括以下三种：

1) unicast单播帧（一对一）：完全相同的MAC地址
2) broadcast广播帧（一对全体）：全1地址
3) multicast多播帧（一对多）：发给局域网部分站点的帧，这种帧需要可编程方法识别，不被所有适配器支持

(2) promiscuous mode混杂方式：听到以太网的帧就直接接收下来，方便网管监视和分析流量，sniffer嗅探器也是用这种方式工作的；但hacker或cracker也能利用这种方式非法获取网上用户的口令

4. DIX Ethernet V2标准的MAC帧的格式

(1) 

(2) 利用以太网的Manchester编码，可以轻松找到数据终结处，FCS又固定为4字节，因此数据段就可定位在14字节~末尾-4字节

(3) 注：为了让MAC长度超过64字节，数据段<46字节时，末尾要填充0

(4) 另外，不使用SONET/SDH的同步传输中，为了让接收端尽快调整好时钟同步频率，第一帧前还需7字节前同步码和1字节帧开始定界符

5. IEEE802.3规定的无效MAC帧：

(1) 帧的长度不是整数个字节；

(2) 帧检验序列 FCS 查出有差错；

(3) 数据字段的长度不在 46 ~ 1500 字节之间，即MAC 帧长度不在 64 ~ 1518 字节之间；

(4) 数据字段的长度与长度字段的值不一致（此处长度字段是IEEE802.3规定里特有的字段，小于0x6000=1536时表示数据段长度，否则表示数据段的类型，由于以太网的垄断，LLC层的消失，这段也没必要了）

1)
2)
3)
4)
5)
6)
7)
8) ----------我是底线----------

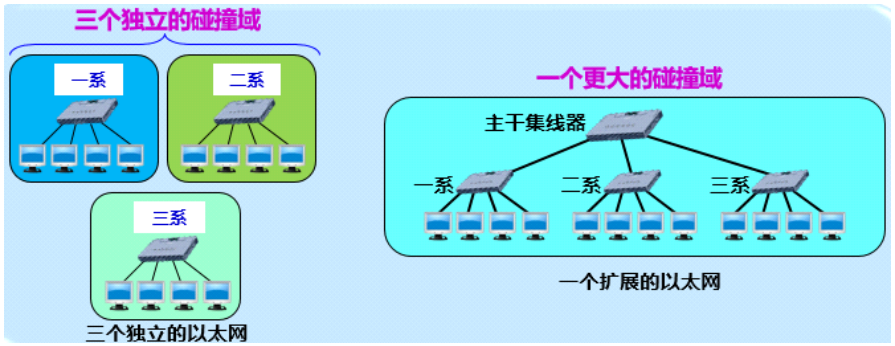
3以太网的发展

2019年5月4日　　15:16

-
- 扩展的以太网

1- 在物理层扩展以太网

1. IEEE802.3规定任两站之间最多可经过三个电缆网段，铜缆时期常使用转发器使距离范围勉强超过数百米
2. 现在扩展主机和集线器距离的简单方法之一就是一对光纤和一对光纤调制解调器了，可以轻松扩展到数千米

(1) 

3. 如上的集线器扩展法不仅使计算机可以轻松“跨系”交流，还实现了以太网地理范围扩展，但也带来了问题：
   (1) 吞吐量无法提高，因为原三个独立collision domain碰撞域合成了一个碰撞域，任一个原碰撞域内想要发送数据，其他两个都不能发送了
   (2) 不同技术的以太网难兼容，如适配器数据率不同时，大家都只能使用最低的数据率，因为**集线器基本只负责转发，不能缓存帧**

2- 在数据链路层扩展以太网

(1) 最初人们使用bridge网桥代替集线器，通过查找地址表，直接转发到对应接口或直接过滤帧（丢弃掉）

(2) 90年switching hub交换式集线器的出现使网桥很快被淘汰，该集线器又称以太网switch交换机或第二层交换机，强调它工作在数据链路层

1. 以太网交换机的特点
   (1) 实质上就是**多接口的网桥**，每个接口连接一台主机或以太网交换机
   (2) 一般工作在全双工通信环境，即不需要CSMA/CD的环境
   (3) 有并行性，能同时连通多对接口，使多对主机同时通信
   (4) 各主机都能独占传输媒体，确保无碰撞地传输（因为以太网交换机每个接口都连接一个独立碰撞域）
   (5) 有存储器，能缓存帧（即存储转发方式），但条件运行时也可直接按MAC地址做cut-through直通转发
   (6) 即插即用，交换表是自学习算法自动建立的
   (7) 专用的交换结构芯片使转发效率高于网桥的查表（不过需要线路速率匹配或协议转换或差错检测时还是会用到软件的）
   (8) 性能远超普通集线器，却不贵（集线器使各用户平分带宽，而交换机相当于

允许各用户独享带宽，并进行累加）

(9) 兼容性好，接入设备一般不需要改动任何软硬件，且交换机内有多重速率的接口，方便各种用户

2. 以太网交换机的自学习

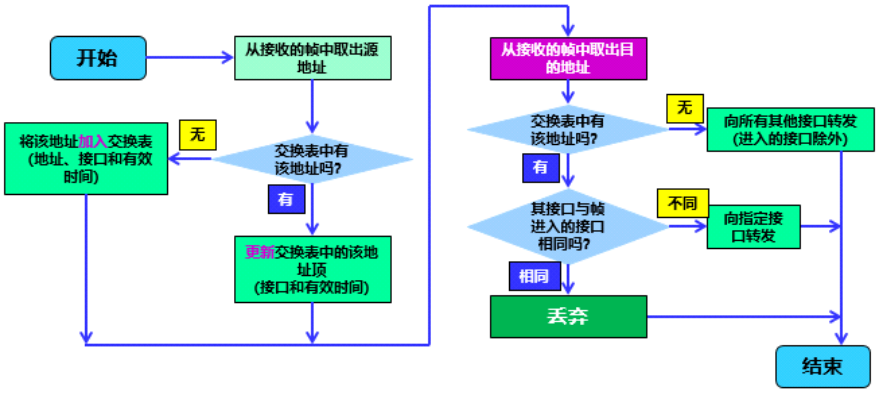
(1) 交换表初始情况是空的

(2) 每次接收到一个帧时，都会把该帧源址和传入接口更新进表

(3) 如果目的地址在表中查不到，则向传入接口以外的接口广播

(4) 隔一段时间会删除表内的项目，防止主机更换接口或适配器

(5)



(6) 为防止两个交换机间的回路被记录在交换表中，IEEE802.1D还制定了Spanning Tree Protocol生成树协议STP来切断逻辑回路

3. 从总线型以太网到星形以太网

(1) 电话网就是星形结构，以太网最初以总线形式出现是因为早期没有研究出像电话交换机一样使用的以太网交换机

(2) 以太网交换机使用全双工方式工作，不需要CSMA/CD协议了，但由于帧结构未改变，还是称为以太网

3- Virtual LAN虚拟局域网

1. IEEE802.1Q标准中的定义：由一些局域网网段构成的与物理位置无关的逻辑组，这些网段有某些共同的需求，每个VLAN的帧都有一个明确的标识符，指明发送帧的源点在哪个VLAN

(1) 实质上是局域网提供给用户的一种服务，不是一种新网

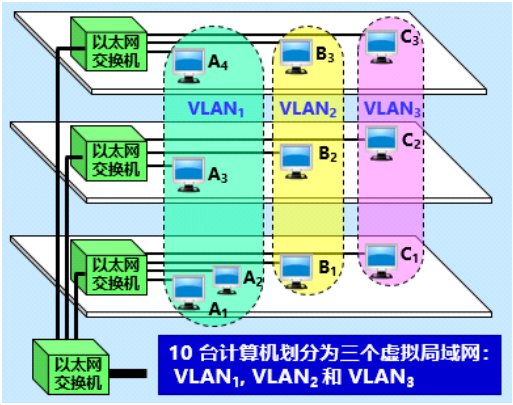
(2) 是用户和网络资源的逻辑组合，可按需将设备和资源重组，方便用户从不同服务器或数据库中获取所需资源

2. broadcast domain广播域：任何一台设备发出的广播通信都能被该部分网络中的所有其他设备所接收的一部分网络

(1) 广播风暴：网络中广播信息过多，引起性能恶化

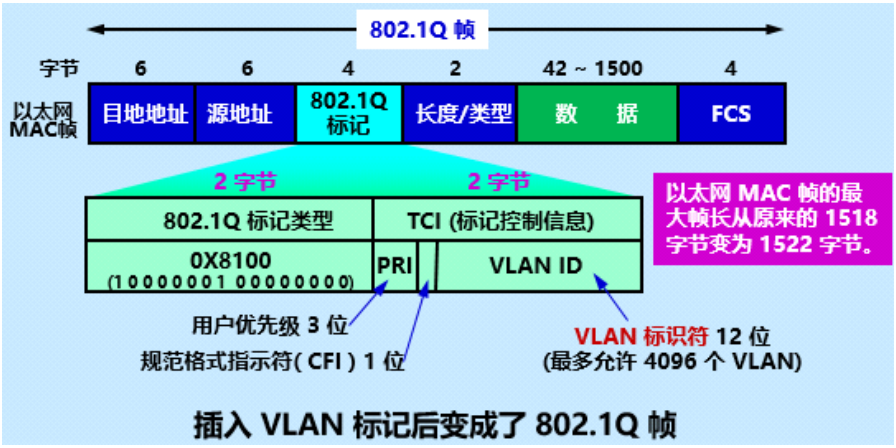
(2) 每个VLAN可视作一个广播域

(3)



3. 虚拟局域网优点：性能改善、管理简化、成本降低、安全性改善
4. 实现方法：
   (1) 基于交换机端口：在物理层实现，简单易用；用户机不能移动
   (2) 基于计算机网卡的MAC地址：在数据链路层实现，用户机可以移动，但增加了管理员的工作量
   (3) 基于协议类型：即根据帧内以太网层类型字段区分，类似上一条
   (4) 基于IP子网地址：根据帧内类型字段和源IP地址字段区分，属于在网络层划分的方法
   (5) 基于高层应用或服务：基于更高层的应用或服务划分，灵活但复杂
5. IEEE802.3ac 标准定义了以太网的帧格式的扩展，以支持虚拟局域网
   (1) 通过在帧内插入4字节的VLAN tag来指明属于哪个VLAN

(2)



-
- 高速以太网

1. 速率>=100Mb/s的以太网即为高速以太网

1- 100BASE-T 以太网/Fast Ethernet快速以太网

1. 是在无屏蔽双绞线上传送 100 Mbit/s 基带信号的星形拓扑以太网
2. 仍使用 IEEE 802.3 的帧格式和 CSMA/CD 协议（即使在全双工环境）
3. 1995 年成为IEEE正式标准，代号 IEEE 802.3u
4. 减小参数a的方法：最大电缆长度缩为100m，最短帧间时隔缩为0.96μs，限制最短帧长为64字节

2- 吉比特以太网

1. 于1996年问世，98年成为IEEE802.3标准，逐渐成为主流产品
2. 特点：
   (1) 允许在 1 Gbit/s 下以全双工和半双工两种方式工作
   (2) 使用 IEEE 802.3 协议规定的帧格式
   (3) 在半双工方式下使用 CSMA/CD 协议，全双工方式不使用
   (4) 与 10BASE-T 和 100BASE-T 技术向后兼容

3. 物理层可使用已有以太网，也可使用ANSI指定的Fiber Channel光纤通道FC
4. 减小参数a的方法：限制铜缆最大长度为100m，使用carrier extension载波延伸，即限制帧长最小512字节，争用期也改为512字节时间
   (1) packet bursting分组突发：想连续发送很多短帧时，只在第一帧载波延伸填充，之后的短帧可一个接一个直接发送，只保留帧间间隔即可
   (2) 对有效payload**载荷**（即数据报部分）无影响，接收端会把延伸的特殊数据删除的
   (3) 全双工环境就不需要载波延伸和分组突发了

3- 10吉比特以太网10GE及更快的以太网（即10^10b）
1. 帧格式与上述以太网完全相同，也保留了 802.3 标准规定的以太网最小和最大帧长，便于升级，不再允许铜缆，不再允许半双工环境，也不再使用CSMA/CD协议
2. 至此，以太网工作范围已经从局域网可以扩展到城域网甚至广域网，并实现了端到端的以太网传输，其优点是
   (1) 实践证明技术成熟：可扩展、灵活、已安装、稳健
   (2) 互操作性很好，各种厂商的以太网都能可靠互操作
   (3) 广域网中使用以太网比SONET和ATM便宜几倍，兼容更多传输媒体
   (4) 统一的帧格式简化了操作和管理，只在帧中继或ATM网络中需要相应接口来互连

4- 用以太网进行宽带接入
1. IEEE 在 2001 年初成立了 802.3 EFM（Ethernet in the First Mile）工作组，专门研究高速以太网的宽带接入技术问题
2. 以太网宽带接入的特点
   (1) 可以提供双向的宽带通信
   (2) 可以根据用户对带宽的需求灵活地进行带宽升级
   (3) 可以实现端到端的以太网传输，中间**不需要再进行帧格式的转换**。这就提高了数据的传输效率且降低了传输的成本
   (4) 可惜地址字段没有用户名字段，即不支持用户身份鉴别
3. PPP over Ethernet在以太网上运行PPP（PPPoE）：将 PPP 帧再封装到以太网中来传输
   (1) 现在的 FTTx 都是使用 PPPoE 的方式进行接入。在 PPPoE 弹出的窗口中键入在网络运营商购买的用户名和密码，就可以进行宽带上网了
   (2) 个人电脑到第一个以太网交换机的带宽是独立的，有保证的，再交到上一级交换机的带宽则是多用户共享的了，因此用户多时网速还是可能降低，需要网络运营商扩容
   (3) 利用 ADSL 进行宽带上网时，从用户个人电脑到家中的 ADSL 调制解调器之间，也是使用 RJ-45 和 5 类线（即以太网使用的网线）进行连接的，并且也是使用 PPPoE 弹出的窗口进行拨号连接的
   (4) 以太网帧经过ADSL调制解调器后即可转换成ADSL使用的PPP帧
      1)
      2)

3)
4)
5) ----------我是底线----------

4网络层

2019年5月4日 18:05

-
- 网络层提供的两种服务

1- 面向连接的Virtual Circuit虚电路

1. 可靠交付由网络负责实现
2. 沿着（逻辑上的）虚电路发送分组即可，首部只需填写虚电路编号
3. 传输完成后释放占用的虚电路
4. 减少了分组的开销，易实现不丢失不重复
5. OSI体系层推出过关于网络层可靠传输服务的虚电路服务标准建议，已成历史

2- 面向无连接的datagram数据报服务

1. 可靠交付由主机实现
2. 互联网的设计思路：网络层只需向上提供best effort delivery尽最大努力交付的、简单灵活、无连接的数据报服务（后来数据报常被称为“分组”）
3. 不需要建立连接，每个分组独立发送
4. 网络层不承诺服务质量，即出错、丢失、重复、失序都可能出现
5. 由主机中的运输层负责差错处理、流量控制等
6. 网络造价大大降低，运行方式灵活，兼容各种应用
7. 如今的互联网证明了这种思路的正确性

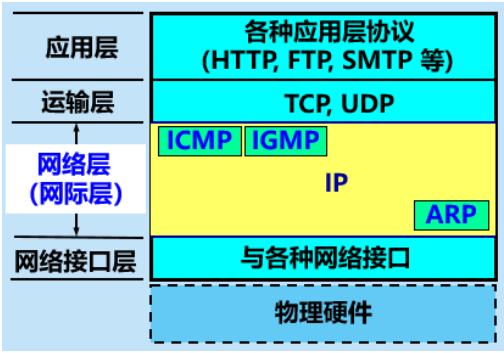
3-

对比的方面	虚电路服务	数据报服务
思路	可靠通信应当由网络来保证	可靠通信应当由用户主机来保证
连接的建立	必须有	不需要
终点地址	仅在连接建立阶段使用，每个分组使用短的虚电路号	每个分组都有终点的完整地址
分组的转发	属于同一条虚电路的分组均按照同一路由进行转发	每个分组独立选择路由进行转发
当结点出故障时	所有通过出故障的结点的虚电路均不能工作	出故障的结点可能会丢失分组，一些路由可能会发生变化
分组的顺序	总是按发送顺序到达终点	到达终点时不一定按发送顺序
端到端的差错处理和流量控制	可以由网络负责，也可以由用户主机负责	由用户主机负责

-
- Internet Protocol网际协议IP

1. 是Robert Kahn和Vint Cerf共同研发的，二人于2005获得图灵奖
2. 除了IPv6外一般都默认讨论的是IPv4因为其他几个版本都没被使用
3. 常与IP协议配套使用的三大协议
   (1) Address Resolution Protocol地址解析协议ARP
   (2) Internet Control Message Protocol网际控制报文协议ICMP
   (3) Internet Group Management Protocol网际组管理协议IGMP
   ☑(4) 其实本来还有Reverse ARP逆解析RARP，DHCP出现后被淘汰了
4. TCP/IP协议族主要关系（上方协议需要使用下方协议）

(1)

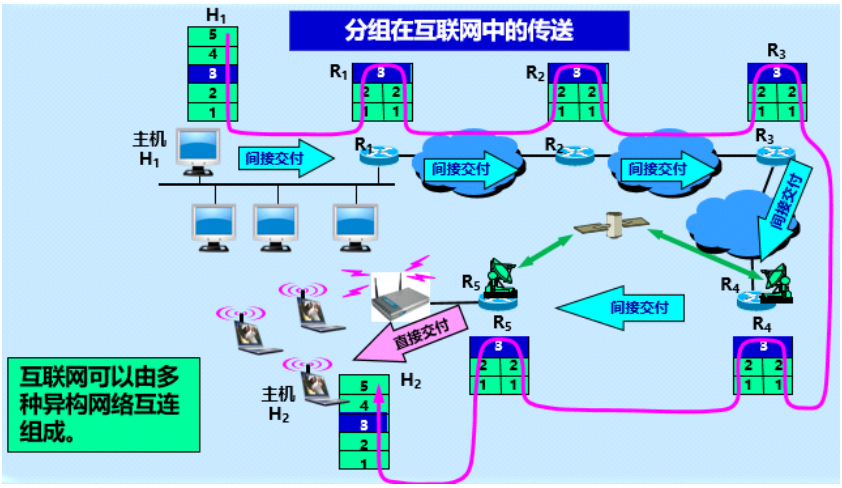


5. 使用了IP协议的network layer网络层也被称为internet网际层或IP层

1- 虚拟互连网络

1. 将网络互连并互相通信，会遇到许多问题需要解决，如：不同的寻址方案；不同的最大分组长度；不同的网络接入机制；不同的超时控制；不同的差错恢复方法；不同的状态报告方法；不同的路由选择技术；不同的用户接入控制；不同的服务（面向连接服务和无连接服务）；不同的管理与控制方式等
2. 将网络互连的中间设备/中间系统/relay中继系统
   (1) 物理层的repeater转发器：仅扩大网络
   (2) 数据链路层的bridge网桥 或 桥接器：扩大网络顺便确认传到哪个接口
   (3) 网络层的router路由器：真正能路由选择地实现不同网络间的互连
      1) 网桥和路由器的混合物：brouter桥路器
   (4) 网络层以上的gateway网关：连接两个不兼容的系统，需要事先在高层做协议转换，较复杂用得较少
   (5) 转发器和网桥连接后仍视为同一个网络，路由器和网关连接的才真正视为网络互连，路由器旧译也为网关
3. 虚拟互连网络/逻辑互联网络/IP网：利用IP协议将物理异构性的网络连成用户视角的统一网络
   (1) 好处：通信时不用管网络内异构的各种细节
   (2) 覆盖全球的IP网上使用TCP协议，即为现在所说的Internet互联网

(3)



2- 分类的IP地址

1. IP 地址及其表示方法，详见RFC791
   (1) IP地址：给每台主机或路由器的每个接口分配一个世界范围内唯一的32位标识符，作用是方便寻址，分配者是Internet Corporation for Assigned Names and Numbers互联网名字和数字分配机构ICANN

(2) 编址方法三阶段

1) 分5类的IP地址，81年就通过了相应的标准协议

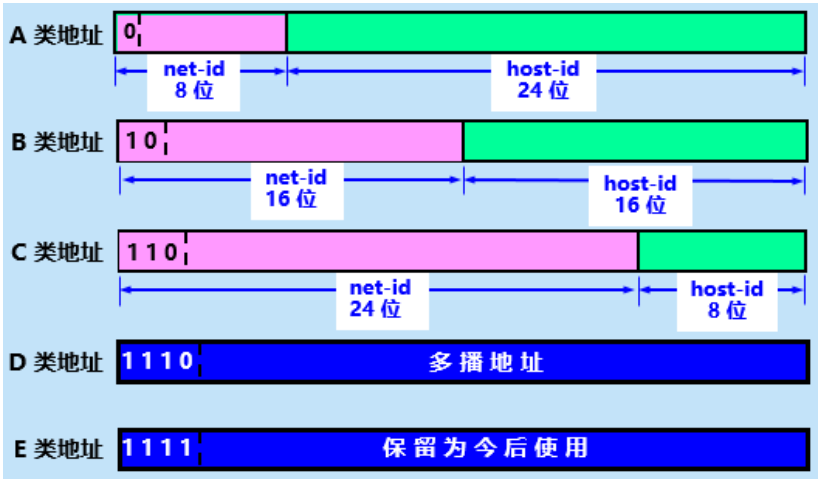
2) 子网划分，改进后的编址方法，85年通过

3) 构成超网，93年后才得到推广

(3) 最初的5类IP地址，前半段是net-id网络号，后半段是host-id主机号

1) IP 地址 ::= { <网络号>, <主机号>}（::==代表“定义为”）

2)



3) ABC类都是单播地址（一对一通信）

4) 现在的路由选择基本都是使用无分类IP地址做路由选择了，ABC类地址已成为历史

(4) dotted decimal notation点分十进制记法：每八位之间加一个点，并转化为十进制（其实相当于加了点的四位256进制数）

2. 常用的三种类别的 IP 地址

(1)

一般不使用的特殊的 IP 地址

网络号	主机号	源地址使用	目的地址使用	代表的意思
0	0	可以	不可	在本网络上的本主机（见 6.6 节 DHCP 协议）
0	host-id	可以	不可	在本网络上的某台主机 host-id
全 1	全 1	不可	可以	只在本网络上进行广播（各路由器均不转发）
net-id	全 1	不可	可以	对 net-id 上的所有主机进行广播
127	非全 0 或全 1 的任何数	可以	可以	用于本地软件环回测试

(2) 主机号全0表示本主机所连接到的单个网络地址，主机号全1表示本主机所在网络的全部主机，所以允许的主机数是2^n - 2，此外：

(3) A类地址中：网络号0x00=0表示this本网络，不可指派（详见6.6节DHCP），网络号0x7f=127表示loopback test环回测试，也不可指派

(4) B类地址中：网络号0x8000=128.0不可指派

(5) C类地址中：网络号0xC00000=192.0.0不可指派

1) （128=10000000，129=10000001，192=11000000）

2) （7=0111，8=1000，C=1100，F=1111）

网络类别	最大可指派的网络数	每个网络中最大主机数	第一个可指派的网络号	最后一个可指派的网络号

(6)

A	**126 (2^7 – 2)**	**16777214**	**1**	**126**
B	**16383 (2^{14} – 1)**	**65534**	**128.1**	**191.255**
C	**2097151 (2^{21} – 1)**	**254**	**192.0.1**	**223.255.255**

3. IP地址的特点

(1) 是一种分等级的地址结构，好处是：

1) IP 地址管理机构在分配 IP 地址时只分配网络号，而剩下的主机号则由得到该网络号的单位自行分配。方便了 IP 地址的管理

2) 路由器仅根据目的主机所连接的网络号来转发分组（而不考虑目的主机号），这使路由表中的项目数大幅度减少

(2) 是标志一个主机（或路由器）和一条链路的接口

1) 当一个主机同时连接到两个网络上时，就必须同时具有两个相应的 IP 地址，其网络号 net-id 必须是不同的

2) 这种主机称为multihomed host多归属主机

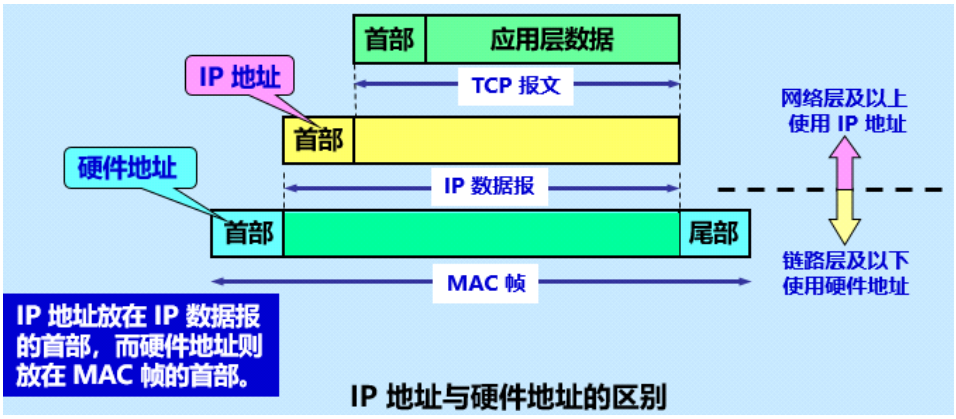
3) 由于一个路由器至少应当连接到两个网络，因此一个路由器至少应当有两个不同的 IP 地址（如果两个路由器直接相连的话，倒是可以不分配 IP 地址，因为这段“网络”只有一条线，称其为unnumbered network无编号网或anonymous无名网）

(3) 用转发器或网桥连接起来的若干个局域网仍为一个网络，具有同样的网络号 net-id

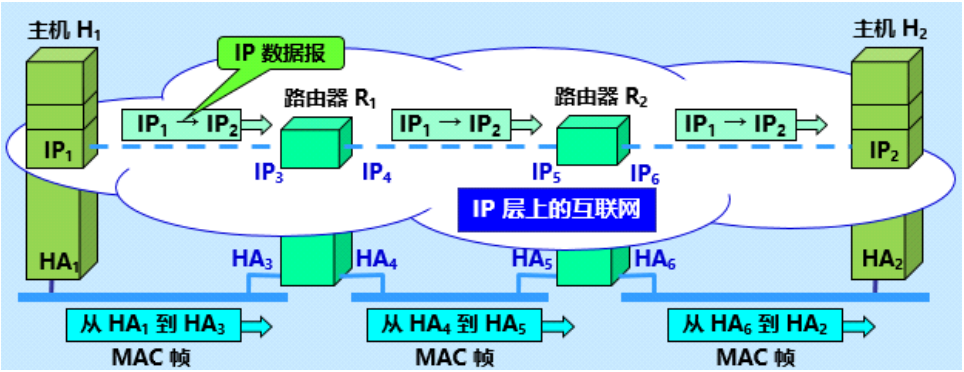
(4) 所有分配到网络号 net-id 的网络，无论范围大小，都是平等的

3- IP地址与硬件地址（逻辑地址与物理地址）

1.



2.



(1) 在 IP 层抽象的互联网上只能看到 IP 数据报

(2) 路由器只根据目的站的 IP 地址的网络号进行路由选择

(3) 在物理网络的链路层只能看见 MAC 帧而看不见 IP 数据报

(4) IP 层抽象的互联网屏蔽了下层的复杂细节，在抽象的网络层上讨论问题，就能够使用统一的、抽象的 IP 地址研究主机和主机或主机和路由器间的通信

(5) **主机 H1 与 H2 通信中使用的IP地址 与 硬件地址 HA**

	在网络层 写入 IP 数据报首部的地址		在数据链路层 写入 MAC 帧首部的地址	
	源地址	目的地址	源地址	目的地址
从 H_1 到 R_1	IP_1	IP_2	HA_1	HA_3
从 R_1 到 R_2	IP_1	IP_2	HA_4	HA_5
从 R_2 到 H_2	IP_1	IP_2	HA_6	HA_2

(6) 路由表的详细得出方式见4.5

(7) MAC首部的MAC目的地址由以下协议得出：

4- Address Resolution Protocol地址解析协议ARP

(1) 不管网络层什么协议，在链路上传数据帧总是要靠硬件地址的

1. 是根据IP地址求出MAC地址的协议，也可以视作数据链路层的协议
2. 局域网各主机或路由器的IP地址到硬件地址的映射表

(1) 存在每个主机的ARP cache高速缓存，里面内容是：

(2) < IP address；MAC address；TTL >

(3) 其中Time To Live是地址映射有效时间，一般为十数分钟

3. 工作步骤：

(1) 能查到目的IP对应的MAC地址就直接发

(2) 查不到的话由ARP进程向局域网上广播发送一个请求分组，内容大意是：IP源址x硬件地址xx请求IPxxx对应的MAC地址

(3) 其他主机的ARP进程都会受到该请求分组

(4) 对应主机的ARP进程单播发送一个响应分组

(5) 双方各自在表内添加对方的新表项，避免短时间内反复广播的通信量

(6) 以上从IP地址到硬件地址的解析是自动进行的，主机用户看不到过程

4. 典型的四种使用情况

(1) 主机想找同网的另一主机，求其硬件地址

(2) 主机想找不同网的另一主机，求最近路由器的硬件地址

(3) 路由器想找同网主机，求其硬件地址

(4) 路由器想找不同网主机，求下一跳路由器的硬件地址

5. 逻辑地址到物理地址的转换过程似乎很复杂，为什么一定要用IP地址呢

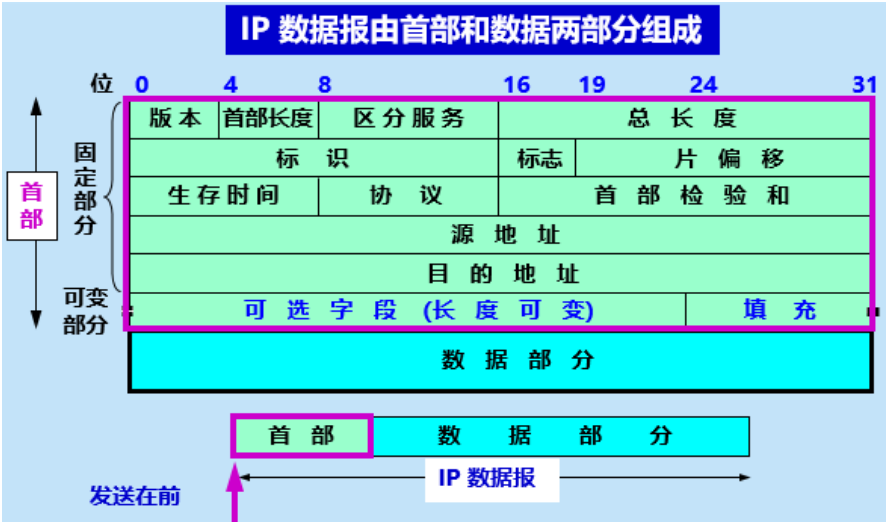
(1) 世界各地的网络使用各种各样的硬件地址，由硬件或软件完成异构网络通信的硬件地址转换是几乎不可能的

(2) 而IP地址使各种主机仿佛连接在同一个网络上一样，且ARP是软件自动进行的，给用户带来了方便

5- IP数据报的格式

1. 格式：首部+数据部分

(1)



2. 首部固定部分

(1) 版本4位（取值4或6）：IP协议的版本

(2) 首部长度4位（取值5~15）：首部位数是32的几倍

(3) 区分服务8位：Differentiated Services，详见8.4.4

(4) 总长度16位：首部+数据的总长度是几字节，超出数据链路层的Maximum Transfer Unit时会被分片，然后会改成分片后的长度

(5) 标识16位：方便分片后再拼回去，这个标识由IP软件提供

(6) 标志3位：相当于三个逻辑变量，最左的暂时没用

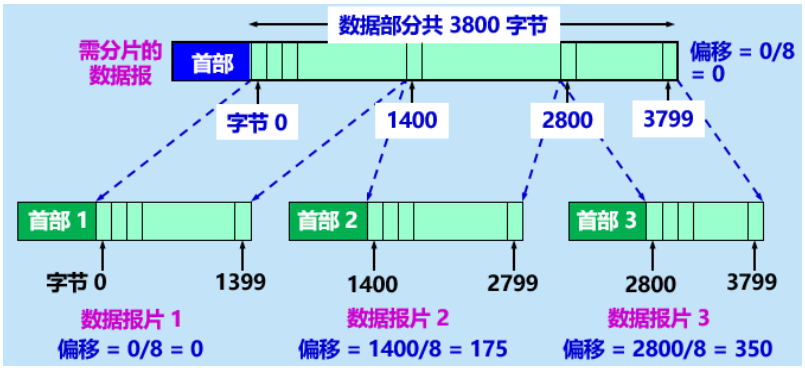
1) 最低位为More Fragment，为0时表示这是最后一片分片

2) 中间位为Don't Fragment，为1时表示不允许分片

(7) 片偏移13位：指出分片在原分组中的位置是64位的几倍处

1) 也就是说每个切片的起点位必须在64的整数倍处

2)



3)

	总长度	标识	MF	DF	片偏移
原始数据报	3820	12345	0	0	0
数据报片1	1420	12345	1	0	0
数据报片2	1420	12345	1	0	175
数据报片3	1020	12345	0	0	350

(8) 生存时间TTL 8 位：减为0后视作交付失败，直接丢弃，现在的实际功能不是以时间为单位，而是以转发过的跳数，即转发一次减一个1

(9) 协议8位：指出数据报的协议内容，方便目的主机的IP层决定上交方式

1) 具体的可以在www.iana.org查到

(10) 首部检验和16位：将除检验和以外的首部分成多个16位字，用异或求和（最高位进1要进到最低位），把和的反码填入这个字段；接收方把包括检验和的

首部分成多个16位的字，求和，得全1就认为无误

(11) 最后是源地址和目的地址各32位

3. 首部可变部分

(1) 是一个选项字段，长度1~40字节，可用于排错、测量、安全

(2) 注意字段内容不足32位整数倍时，末尾要填充0

(3) 可变长度部分的存在增加了路由器处理数据报的开销，于是v6固定了首部的长度

6- IP层转发分组的流程

1. 分组转发算法

(1) 从数据报的首部提取目的主机的 IP 地址 D, 得出目的网络地址为 N

(2) 若网络 N 与此路由器直接相连，则把数据报直接交付目的主机 D；否则是间接交付

(3) 若路由表中有目的地址为 D 的特定主机路由，则把数据报传送给路由表中所指明的下一跳路由器

(4) 若路由表中有到达网络 N 的路由，则把数据报传送给路由表指明的下一跳路由器

(5) 若路由表中有一个默认路由，则把数据报传送给路由表中所指明的默认路由器

(6) 报告转发分组出错

2. 路由表的内容：

(1) 每行对应一个网络，第一列是目的网络IP（末尾全0）

(2) 每行第二列是下一跳地址，对应到这个网络的拓扑结构的下一结点

(3) 下一跳为另一个地址时表示需要间接交付给别的路由器；为全零时表示本路由器与目的主机在同一网络，可以直接交付

(4) 偶尔会发现第一列不是末尾全零的网络而是指定主机的IP地址，这是网管为了方便控制或测试网络或考虑安全问题时设的特定主机路由

3. default route默认路由行的第一列是全0

(1) 路由器对外连接较少时甚至可以直接不搜路由表，直接交给默认路由

4. 传送给下一跳的意思：由网络接口软件按ARP协议，将下一跳IP地址转为MAC地址，再交给数据链路层作为首部，由数据链路层转发给它

1)

2)

3)

4)

5)

6)

7)

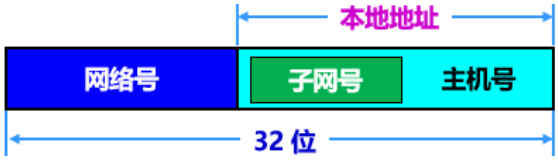
8) ----------我是底线----------

4划分子网和构造超网和ICMP协议

2019年5月4日 22:58

-
- 划分子网和构造超网

1- 划分子网

1. 从两级IP地址到三级IP地址（1985年的RFC950）
   (1) 两级IP地址该批判的地方
      1) 空间利用率低：指主机号一般用不完
      2) 路由表太大影响性能：因为每个物理网络都要分配一个网络号
      3) 两级不够灵活：指申请新网络一定要申请新网络号
   (2) subnet划分子网/子网寻址/子网路由选择是85年起新增的地址字段
      1) 单位内可将物理网络划分为若干子网
      2) 但对单位外仍表现为一个网络，即对外网透明
      3) 划分方法是从主机号借几位作为subnet-id子网号
      4) 发给该单位的IP数据包仍是交给该单位的路由器，由其找到对应子网并交付给目的主机
      5) RFC950规定子网号不能为全0或全1，不过CIDR的广泛使用使全1或全0的子网号也可能出现了，但部分路由器可能暂时不支持
      6) 

      7) IP地址 ::= {<网络号>, <子网号>, <主机号>}
   (3) 优点
      1) 减少了IP地址的浪费
      2) 使网络的组织更加灵活
      3) 更便于维护和管理
   (4) 缺点：每个网络上可连接的主机总数会减少
2. subnet mask子网掩码
   (1) 用于找出IP地址的子网部分，长度也和IP地址长度一样是32位
   (2) 左边一连串1的位数=网络号+子网号的位数
   (3) 右边一连串0的位数=主机号的位数
   (4) 使用方法：与IP地址诸位做逻辑与运算
   (5) 三类常用地址的默认掩码的1的位数=网络号的位数，即没有子网
3. 子网掩码的性质
   (1) 是网络或子网的重要属性
   (2) 路由器间交换路由信息时必须传递子网掩码
   (3) 路由器每个表项/每行都要有目的网络地址及对应子网掩码
   (4) 同时连接在两个子网的路由器拥有两个网络地址和两个子网掩码
4. 子网划分方法

(1) 定长子网掩码划分：每个子网的掩码都相同
(2) Variable Length Subnet Mask变长子网掩码VLSM划分：子网掩码可能不同，子网大小也不同

2- 使用子网时分组的转发

1. 有子网时，子网掩码会影响寻址，而数据报本身并不会提供子网掩码
   (1) 由支持使用子网划分的路由表提供子网掩码，即路由表每个表项应当有三个内容：目的网络地址、子网掩码、下一跳地址
   (2) 新算法大致是**先判断是否在同一子网**，能否直接交付，不能的话交给路由器，先找可直接交付的子网，再找特定主机路由，再找有下一跳的子网，再找默认路由，关于路由器的算法如下：
2. 有子网的路由器转发分组算法
   (1) 从收到的分组的首部提取目的 IP 地址 D
   (2) 先用各网络的子网掩码和 D 逐位相“与”，看是否和相应的网络地址匹配。若匹配，则将分组直接交付。否则就是间接交付，执行(3)
   (3) 若路由表中有目的地址为 D 的特定主机路由，则将分组传送给指明的下一跳路由器；否则，执行 (4)
   (4) 对路由表中的每一行，将子网掩码和 D 逐位相“与”。若结果与该行的目的网络地址匹配，则将分组传送给该行指明的下一跳路由器；否则，执行 (5)
   (5) 若路由表中有一个默认路由，则将分组传送给路由表中所指明的默认路由器；否则，执行 (6)。
   (6) 报告转发分组出错

3- 构造超网和Classless Inter-Domain Routing无分类域间路由选择CIDR

(1) 92年互联网危机：
   1) B类地址已经分配完一半了
   2) 路由表项目数增长到了几万
   3) IPv4地址终将耗尽（最终IANA在11年正式宣布了v4地址耗尽）
(2) 于是94年IETF开始尝试无分类编址来解决前两个问题，并开始计划IPv6来解决第3个问题
(3) 于是在87年的RFC1009中的变长掩码VLSM基础上研究出了CIDR
(4) CIDR地址中消除了三类地址和划分子网的概念，更有效地分配了IPv4的地址，使IPv6投入使用前允许互联网规模继续增长

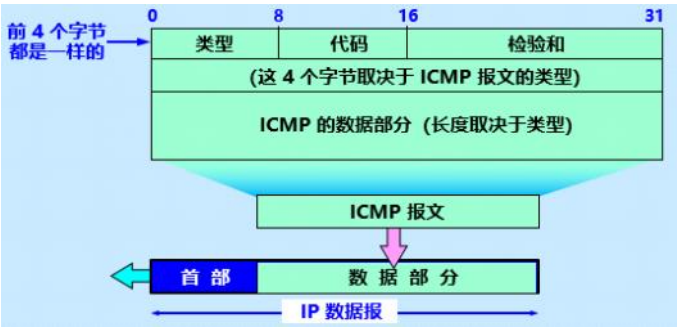
1. network-prefix网络前缀
   (1) CIDR的IP地址 ::= {<网络前缀>, <主机号>}
   (2) slash notation斜线记法或称CIDR记法：在IP地址后写'/'和前缀位数
   (3) 简写法：点分十进制中低位的连续0可省略
      1) 例：10.0.0.0/10 可简写为 10/10
   (4) 星号分割法：在前缀和主机号之间加一个星号
2. address mask地址掩码：32位，用法同子网掩码，可视作广义的子网掩码
   (1) 地址掩码中1的位数即为前缀的位数
   (2) 地址掩码中0的位数即主机号的位数

3. CIDR地址块：由网络前缀都相同的连续的 IP 地址组成的集合
   (1) 注意全0和全1的主机号地址一般不使用
   (2) 例：128.14.32.0/20 地址块的最小地址或称起始地址：128.14.32.0；最大地址：128.14.47.255
   (3) 算上全1和全0的话，地址块内的地址数一定有2的整数次幂个
   (4) 理想的地址块分配是按地理位置分的，能加大路由速度，可惜在CIDR投入使用前的地址管理机构没有想到这一点，不过CIDR已经尽力推迟了地址耗尽
4. supernetting构成超网：将同一CIDR地址块的地址合并到一个表项
   (1) 减少了路由器间路由信息的交换，从而提高互联网的性能
   (2) 又称route aggregation路由聚合
   (3) 称为超网的原因：前缀长度<23的CIDR地址块都包含了多个C类地址
   (4)

CIDR 前缀长度	点分十进制	包含的地址数	相当于包含分类的网络数
/13	255.248.0.0	512 K	8 个 B类或 2048 个 C 类
/14	255.252.0.0	256 K	4 个 B 类或1024 个 C 类
/15	255.254.0.0	128 K	2 个 B 类或512 个 C 类
/16	255.255.0.0	64 K	1 个 B 类或256 个 C 类
/17	255.255.128.0	32 K	128 个 C 类
/18	255.255.192.0	16 K	64 个 C 类
/19	255.255.224.0	8 K	32 个 C 类
/20	255.255.240.0	4 K	16 个 C 类
/21	255.255.248.0	2 K	8 个 C 类
/22	255.255.252.0	1 K	4 个 C 类
/23	255.255.254.0	512	2 个 C 类
/24	255.255.255.0	256	1 个 C 类
/25	255.255.255.128	128	1/4 个 C 类
/26	255.255.255.192	64	1/4 个 C 类
/27	255.255.255.224	32	1/8 个 C 类

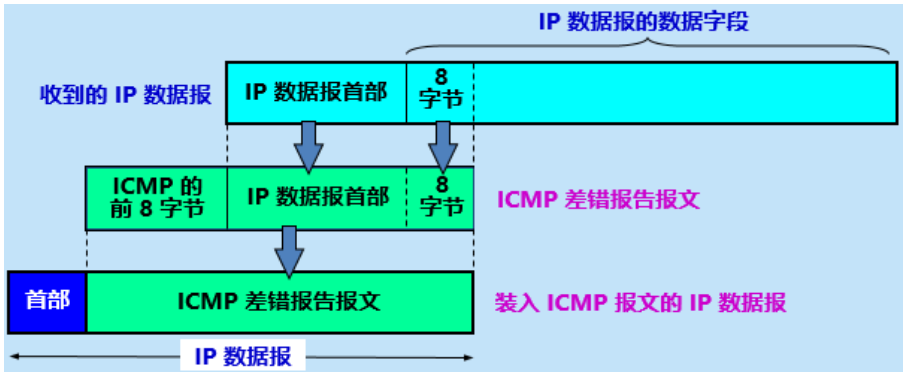
5. longest-prefix matching最长前缀匹配：为了路由到更specific具体的的地址，路由表匹配结果中优先选网络前缀最长的，称为最长匹配或最佳匹配
6. binary trie二叉线索查找路由表
   (1) 将每个IP地址的unique prefix唯一前缀存在二叉树中
   (2) 左子的边对应0，右子的边对应1，则向下的路径即为地址
   (3) 每按层次向下一次就尝试匹配一次
   (4) 压缩技术：如共同前缀可压缩在一条边内
      -
      - 网际控制报文协议ICMP

1. Internet Control Message Protocol是为了更有效地转发 IP 数据报和提高交付成功的机会，而在网际层使用的协议
2. 是互联网的标准协议

3. 允许主机或路由器报告差错情况和提供有关异常情况的报告
4. 是IP数据报中的、IP层的协议
5. 

4- 报文种类

1. ICMP差错报告：终点不可达、时间超过、参数问题、Redirect改变路由

(1) 

(2) 不应发送差错报告的情况：对 ICMP 差错报告报文不发送、对第一个分片的数据报片的所有后续数据报片都不发送、对具有多播地址的数据报不发送、对具有特殊地址（如127.0.0.0 或 0.0.0.0）的数据报不发送

2. ICMP询问：回送请求和回答报文；时间戳请求和回答报文。曾经还用过信息请求与回答报文、掩码地址请求和回答报文、路由器询问和通告报文、源点抑制报文

5- 应用举例

1. Packet InterNet Groper分组网间探测PING
   (1) 用来测试两主机的连通性
   (2) 使用了ICMP回送请求和回答报文
   (3) 是应用层直接使用网络层ICMP的经典例子，没使用TCP或UDP
2. TraceRoute跟踪路径，在windows的cmd里简写为tracert
   (1) 利用TTL分别为1，2，……的一系列由于非法端口而无法交付的UDP数据包，实现让跳数距离递增的各路由器自动发返ICMP时间超过差错报告
      1)
      2)
      3)
      4)
      5) ----------我是底线----------

4路由选择协议

2019年6月16日 15:29

-
- 互联网的路由选择协议

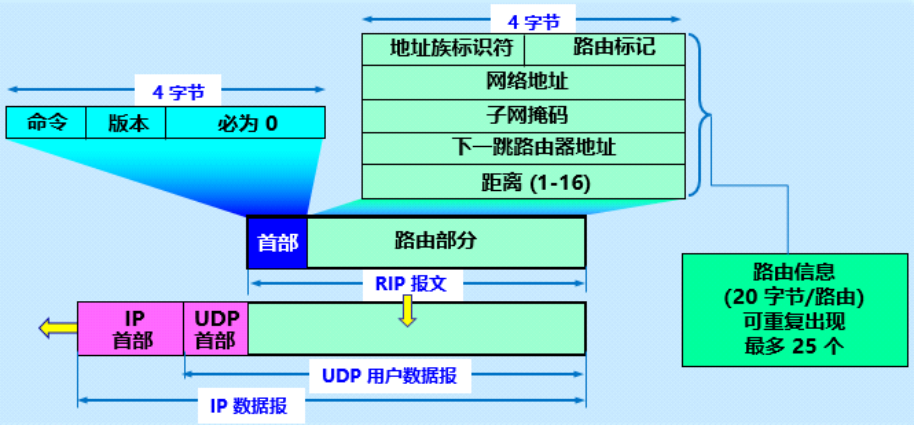
1- 有关路由选择协议的几个基本概念

1. 理想的路由算法：正确完整、计算简单、自适应性（适应网络拓扑变化）或称tobustness稳健性、稳定性（可收敛）、对用户公平、最佳
   (1) 最佳一般指在考虑某一条件下得出的最合理选择
   (2) 路由选择需要网络中各结点共同协调工作，网络环境的变化一般无法事先知道
   (3) 静态路由选择：非自适应性选择，适用于小范围，对应动态自适应性
2. 分层次的路由选择协议
   (1) 互联网规模大，路由表信息有限
   (2) 许多单位不愿将内部网络细节公开，但又需要连到互联网
   (3) Autonomous System自治系统AS：单一技术管理下的一组路由器，使用内部路由协议以确定AS内的路由，AS对其他AS表现为单一的一致的路由选择策略
   (4) Interior Gateway Protocol内部网关协议：AS内使用的协议，负责intradomain routing域内路由（注：此处网关指路由器）
   (5) External Gateway Protocol外部网关协议：AS间的路由协议，负责interdomain routing域间路由（注：早期最常用的EGP就叫EGP）

2- 内部网关协议Routing Information Protocol

1. 工作原理：基于距离（跳数）向量的路由选择协议
   (1) 定时地、仅和相邻的、路由器交换全部的路由表
   (2) 适用于小型网络，经若干次交换后会收敛出连通域内所有结点最短距离
   (3) 最短路converge收敛得很快，但对坏消息（故障）传得很慢
2. 距离向量算法（类似Bellman-Ford或Ford-Fulkerson）
   (1) 把相邻路由器发来的RIP报文的项目的下一跳改为该路由器，对跳数++
   (2) 若出现新表项，或对原表项做出了更新，或出现了比原表项距离更小的项，则在路由表中做出修改
   (3) 若3分钟未收到相邻路由表的新报文，则将该路由器视作不可达，对其距离改为16
3. RIP2报文格式

(1)



(2) RIP2 报文中的路由部分由若干个路由信息组成，每个路由信息需要20字节。地址族标识符（又称为地址类别）字段用来标志所使用的地址协议

(3) 路由标记填入Autonomous System Number自治系统号（由IANA分配，0~65535），因为 RIP 有可能收到本AS外的路由信息

(4) 再后面指出某个网络地址、该网络的子网掩码、下一跳路由器地址以及到此网络的距离

(5) 一个 RIP 报文最多可包括 25 个路由，因而 RIP 报文的最大长度是 4+20 x25=504 字节

(6) 第一个路由信息也可用作鉴别

3- 内部网关协议Open Shortest Path First

1. 特点

(1) 此处open指不受任一厂商控制而公开发表，shortest是考虑权值的短，OSPF2成为了互联网标准协议（RFC2328）

(2) 使用了分布式的link state protocol链路状态协议

1) 仅当链路状态发生变化时才发送更新信息

2) 向AS内所有路由器用flooding洪泛法发送信息，100ms内可以响应网络变化

3) 发送的是本路由器与所有相邻路由器间的链路状态metric度量（不过对多点接入局域网可能采用designated router指定路由器，以减少广播信息量）

(3) 相当于所有路由器都有一个link-state database，知道了全网范围内的拓扑结果图，再通过Dijkstra算法快速收敛更新过程

1) 有相同的数据库即可视为“同步”的或fully adjacent完全邻接

(4) OSPF会将AS再分为许多层次的area

1) 最上层的称为backbone area主干区，其标识符为0.0.0.0

2) 主干区内的路由器都是backbone router主干路由器

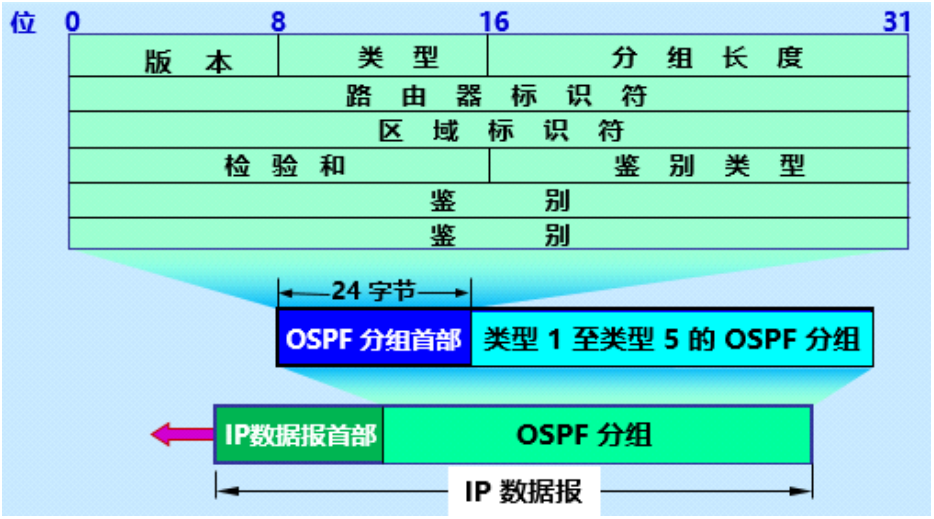
3) 每个区都有32位标识符，最好一个区不要超过200个路由器

4) 其他区域的信息都交给各自的area border router概括

(5) 不使用UDP，直接用IP数据报传送，协议字段为89。数据报很短，一般不会被分片

2. 首部格式

(1)

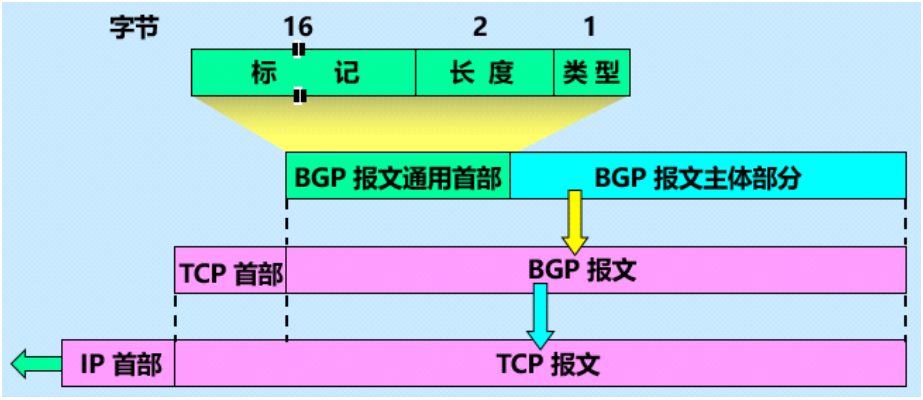


3. 优点
   (1) 对不同类型的业务会计算出不同路由
   (2) 等代价的路径会被平均分配通信量，实现load balancing负载平衡
   (3) 有鉴别功能，保证仅在可信赖的路由器间交换的信息
   (4) 支持可变长子网划分和CIDR
   (5) 每5秒钟，允许链路状态序号++一次，实现600年内不重复的序号比较链路状态的新旧程度
4. 五种分组类型：hello问候（确认可达性）、database description数据库描述（同步数据库）、link state request链路状态请求、link state update链路状态更新（洪泛法更新）、link state acknowledgement链路状态确认

4- 外部网关协议Border Gateway Protocol

1. 是不同AS的路由器间交换路由信息的协议，通过在AS间交换可达性信息，实现跨AS的路由选择，不求最佳路由，只求不出现回路
2. BGP speaker发言人：通过共享网络连接在一起的路由器，一般是BGP边界路由器，但也可以不是
   (1) 因而协议结点数量级与AS相同，这比许多AS内的网络数少
   (2) 发言人数量不会太多，因此AS间路由不会太难
   (3) 支持CIDR，即表项字段为网络前缀、下一跳、AS序列
   (4) 刚运行时与邻站交换整个路由表，之后只更新变化部分，对节省网络带宽和减少路由器的处理开销都有好处
3. 交换路由信息
   (1) BGP发言人间先建立TCP连接，交换BGP报文实现BGP session会话
   (2) 这样的两个发言人彼此成为对方的neighbor邻站或peer对等站
   (3) TCP可靠服务简化了该路由选择协议
   (4) 交换的信息是到其他网络需经过的一系列AS（path vector路径向量）
   (5) 检查自己不在该向量中，则能保证更新到没有回路的路由信息
4. 通用首部

(1)

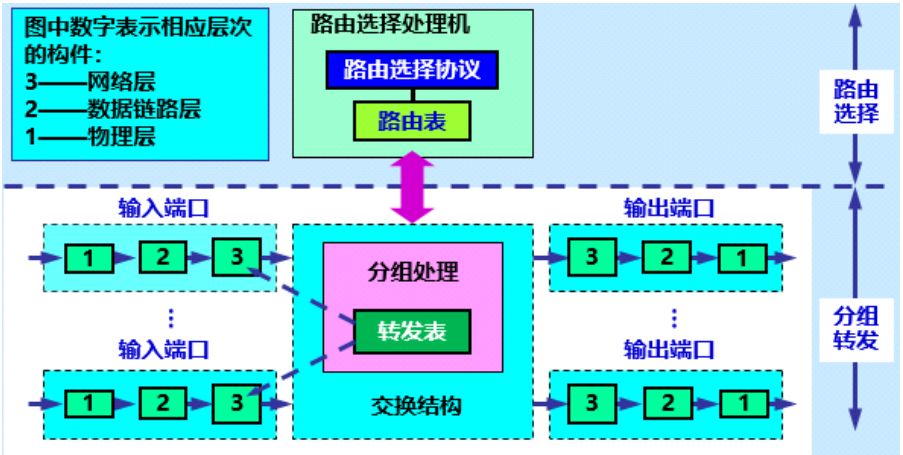


5. 四种报文：
   (1) open建立邻站连接，包括1字节版本号暂时固定为值=4；2字节全球标准AS号；2字节保持邻站关系时间；4字节BGP标识符即IP地址和可选参数长度及可选参数
   (2) update更新路由信息，包括2字节不可行路由长度及待撤销的具体不可行路由；2字节新增路径总长度及具体新增路径；Network Layer Reachability Information网络层可达信息NLRI（该网络的网络号）
   (3) keepalive确认邻站关系，仅含通用首部
   (4) notification发送检测到的差错，包含1字节差错代码；1字节差错子代码；和差错数据

5- 路由器的构成
   (1) 是网络层设备，主要作用是连通不同网络和选择信息传送线路；提高速率，减轻负荷，节约资源，让网络系统发挥更大的效益

1. 物理结构
   (1) 是具有多个输入端口和多个输出端口的专用计算机，任务是转发分组
   (2) 不断挑选合适的端口转发给下一跳，直至到达终点

(3)



   (4) 路由选择部分/控制部分：核心是路由选择处理机，根据协议维护路由表
   (5) 分组转发部分：输入输出端口和switching fabric交换结构/组织，根据forwarding table转发表处理分组

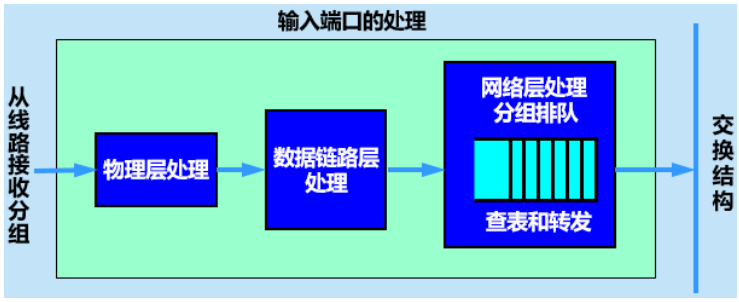
2. forwarding转发和routing路由
   (1) forwarding是按转发表将IP数据包从合适端口转出
   (2) routing是按分布式算法动态选择下一跳

3. 转发输入
   (1) 输入端口里装有前三层的处理模块，去掉数据链路层的首尾后，将分组送到

网络层的队列中排队

(2) 输入端口的查找和转发功能在路由器的交换功能中是最重要的

(3)



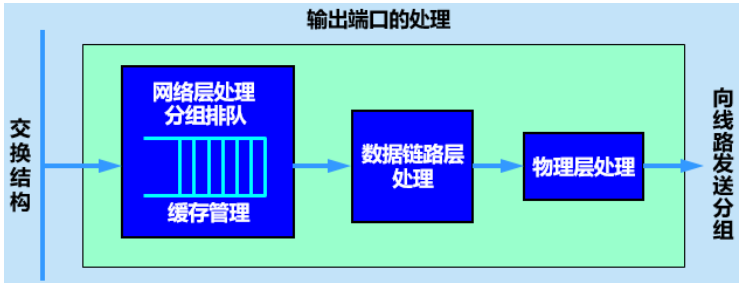
4. 转发输出

(1) 输出端口里也装有前三层的处理模块，负责从交换结构接受分组，再发送到路由器外的线路，交换结构传送的数据较多时暂存进缓冲队列

(2) 数据链路层的处理模块会给分组加上数据链路层的首尾，再交给物理层

(3) 输入或输出时的缓冲队列存储空间满时都将不得不丢弃分组

(4)



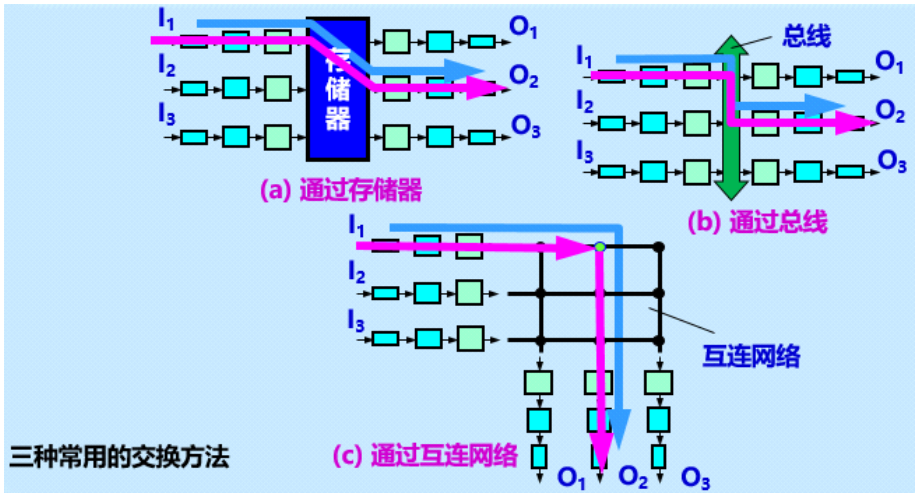
5. 交换结构是关键构件，负责把输入端的分组转移到合适的输出端

(1) 存储器交换法：输入端口用中断通知路由选择处理机按目的地址查路由表，复制到对应输出端口的缓存。交换速率不大于读写速率/2

(2) 总线法：用共享总线直接从输入端口转到输出端口。速率不超过总线速率，但随着总线带宽提高到吉比特每秒，总线型流行了起来

(3) crossbar switch fabric纵横交换结构法/interconnection network互联网络法：N个输入端口对应N个水平总线，N个输出端看对应N个垂直总线，当输出端口对应垂直总线不空闲时将留在输入端口排队

(4)



1)

2)

3)

4)

5)
6)
7)
8) ----------我是底线----------

4IPv6、多播、VPN

2019年6月16日 21:29

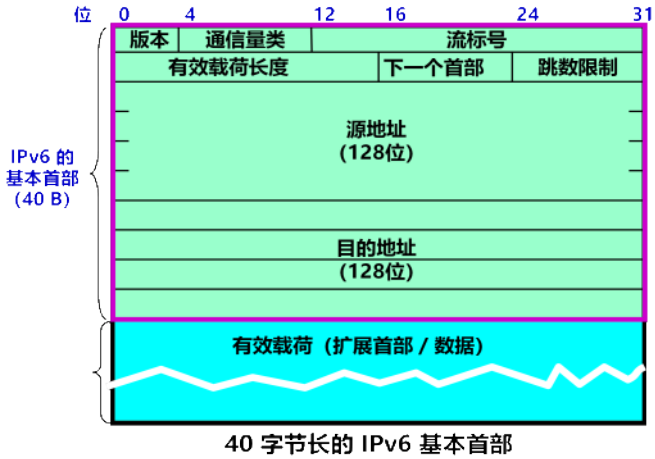
-
- IPv6

1. 到 2011 年 2 月，IPv4 的 32 位地址已经耗尽。
2. ISP 已经不能再申请到新的 IP 地址块了。
3. 我国在 2014 – 2015 年也逐步停止了向新用户和应用分配 IPv4 地址。
4. 解决 IP 地址耗尽的根本措施就是采用具有更大地址空间的新版本的 IP，即 IPv6。

1- IPv6首部

1. IPv6 仍支持无连接的传送，但将协议数据单元 PDU 称为分组。为方便起见，本书仍采用数据报这一名词。所引进的主要变化如下：
   (1) 更大的地址空间：IPv6 将地址从 IPv4 的 32 位 增大到了 128 位。
   (2) 扩展的地址层次结构。
   (3) 灵活的首部格式： IPv6 定义了许多可选的扩展首部。
   (4) 改进的选项： IPv6 允许数据报包含有选项的控制信息，其选项放在有效载荷中。
   (5) 允许协议继续扩充。
   (6) 支持即插即用（即自动配置）：因此 IPv6 不需要使用 DHCP。
   (7) 支持资源的预分配： IPv6 支持实时视像等要求，保证一定的带宽和时延的应用。
   (8) IPv6 首部改为 8 字节对齐：首部长度必须是 8 字节的整数倍（默认40，可在有效载荷中扩展）。原来的 IPv4 首部是 4 字节对齐。
2. 数据报途中经过的路由器都不处理这些扩展首部（只有一个首部例外，即逐跳选项扩展首部），这样就大大提高了路由器的处理效率。

(1)



40 字节长的 IPv6 基本首部

3. 扩展首部：逐跳选项、路由选择、分片、鉴别、封装安全有效载荷、目的站选项

2- IPv6的地址：单播、多播、任播

1. 结点和接口
   (1) IPv6 将实现 IPv6 的主机和路由器均称为结点
   (2) 一个结点就可能有多个与链路相连的接口

(3) IPv6 地址是分配给结点上面的接口的

1) 一个接口可以有多个单播地址

2) 其中的任何一个地址都可以当作到达该结点的目的地址。即一个结点接口的单播地址可用来唯一地标志该结点

2. colon hexadecimal notation冒号十六进制记法

(1) 形如68E6:8C64:FFFF:FFFF:0:1180:960A:FFFF

(2) 允许将一连串连续0用双冒号取代

(3)

地址类型	二进制前缀
未指明地址	**00…0（128位），可记为 ::/128。**
环回地址	**00…1（128位），可记为 ::1/128。**
多播地址	**11111111（8位），可记为 FF00::/8。**
本地链路单播地址	**1111111010（10位），可记为 FE80::/10。**
全球单播地址	**（除上述四种外，所有其他的二进制前缀）**

3. IPv4向v6的过渡

(1) dual stack双协议栈：在完全过渡到 IPv6 之前，使一部分主机（或路由器）装有两个协议栈，一个 IPv4 和一个 IPv6

1) 双协议栈的主机（或路由器）记为 IPv6/IPv4，表明它同时具有两种 IP 地址：一个 IPv6 地址和一个 IPv4 地址

2) 双协议栈主机在和 IPv6 主机通信时是采用 IPv6 地址，而和 IPv4 主机通信时就采用 IPv4 地址

3) 根据 DNS 返回的地址类型可以确定使用 IPv4 地址还是 IPv6 地址

(2) 隧道技术：在 IPv6 数据报要进入 IPv4 网络时，把 IPv6 数据报封装成为 IPv4 数据报，整个的 IPv6 数据报变成了 IPv4 数据报的数据部分；当 IPv4 数据报离开 IPv4 网络中的隧道时，再把数据部分（即原来的 IPv6 数据报）交给主机的 IPv6 协议栈

4. ICMPv6：IPv6 也不保证数据报的可靠交付，因为互联网中的路由器可能会丢弃数据报。因此 IPv6 也需要使用 ICMP 来反馈一些差错信息。新的版本称为 ICMPv6

(1) 地址解析协议 ARP 和网际组管理协议 IGMP 协议的功能都已被合并到 ICMPv6 中

(2) ICMPv6 是面向报文的协议，它利用报文来报告差错，获取信息，探测邻站或管理多播通信。

(3) ICMPv6 还增加了几个定义报文的功能及含义的其他协议

1) ND (Neighbor-Discovery)：邻站发现

2) MLD (Multicast Listener Delivery)：多播听众交付

-
- IP多播

1- IP multicast多播/组播的基本概念

1. IP多播：互联网上进行多播，靠路由器实现

(1) 能运行多播协议的路由器称为multicast router多播路由器
(2) 92年起，Multicast Backbone On the InterNEt多播主干网MBONE开始试验，现在已有相当大的规模了

2. 目的：更好的支持一对多通信（一个源点发到多个终点）
3. 多播IP地址：IP多播传送的分组使用的地址
   (1) 多播数据包的目的地址里写的是多播组的标识符
   (2) IPv4的每个D类地址对应一个多播组
   (3) 多播地址只能用于目的地址，不能是源地址
4. 多播数据报：首部协议字段值为2（IGMP）的数据报
   (1) 尽最大努力交付，不保证交给多播组所有成员
   (2) 不会产生ICMP差错报文，例：PING多播地址不会收到响应

2- 在局域网上硬件多播

1. IANA拥有的以太网地址块高24位为00-00-5E?
   (1) 规定IANA拥有的可用于以太网硬件地址的地址号只有后23位可变
   (2) 规定第8位为1时该地址为多播地址
   (3) 即实际以太网多播地址范围01-00-5E-00-00-00 ~ 01-00-5E-7F-FF-FF
2. d类IP地址前4位固定为1110，规定后28位中的后23位对应上一行的硬件地址
   (1) 因为d类IP地址后28位的前5位与硬件地址无关，因此在将多播IP地址与以太网硬件地址进行映射时可能需要IP层软件过滤

3- 网际组管理协议IGMP和多播组路由选择协议

1. IP多播两种协议
   (1) Internet Group Management Protocol网际组管理协议使路由器能知道多播组成员的信息
      1) 并不知道多播组的成员数，也不知道分布在哪些网络上，只需让多播路由器能知道其局域网上是否有主机加入或退出某多播组
   (2) 多播路由选择协议使多播路由器能互相协同工作、以最小代价传送多播数据包给组成员
      1) 需动态适应多播组成员的进出（不像ICMP只关注拓扑结构变化）
      2) 多播数据报可以由没加入多播组的主机发出，也可从非成员发出
      3) 转发时要考虑从哪来到哪去
2. 网际组管理协议IGMP
   (1) 和ICMP一样利用IP数据包传递报文，也向IP提供服务，一般视作网际协议IP的一个组成部分而不是一个单独的协议
   (2) 02年的rfc3376确立了IGMPv3为建议标准
   (3) 加入多播组：新入主机向该组多播地址发IGMP报文，多播路由器收到后，将新的组成员关系转发给互联网的其他多播路由器
   (4) 探询组成员变化：周期性（默认125秒）探询本地局域网的主机在最长响应时间N秒（默认为10）有没有响应，多次无响应后不将该组的成员关系转发给其他多播路由器
   (5) 主机和多播路由器间所有通信都使用IP多播；一个网络有多个多播路由器

时，只有其中一个会探询；只要有一个回答，其他的就可以不响应探询

3. 多播路由选择：找出以源主机为根的多播转发树
   (1) 洪泛与剪除法：适用于较小多播组，用洪泛广播转发多播数据报
      1) Reverse Path Broadcasting反向路径广播RPB：收到多播数据报时检查它是否是从源点沿最短（跳数）路径传来的，是的话才向其他地方转发，有多条最短路时选择IP地址最小的邻站
      2) 洪泛路径即为该源点的多播转发树
      3) 子结点方向无组成员时应剪除该结点及其子树；有新增组成员时再考虑接入到该多播转发树上
   (2) tunneling隧道技术：适用于地理位置上很分散的多播组
      1) 在进出不支持多播的网络时加上普通数据报的首部，封装成单播
      2) “出隧道”后再重新当作多播数据包，这种封装称为IP in IP
   (3) 基于核心的发现技术：适用于大范围的多播组
      1) 给多播组指定一个core核心路由器，给出其IP单播地址
      2) 发往该组的数据报都发给核心，再根据核心的转发树转发；若在发往核心路上恰经过目标路由器，则提前截获；有主机申请加入该组时，则用隧道技术向其转发各多播数据报的副本
   (4) 其他协议：Distance Vector Multicast Routing Protocol距离向量多播路由选择协议 DVMRP（RFC1075）、Core Based Tree基于核心的转发树 CBT（RFC2189，2201）、Multicast Extensions to OSPF开放最短通路优先的多播扩展 MOSPF（RFC1585）、Protocol Independent Multicast-Sparse Mode协议无关多播-稀疏方式 PIM-SM（RFC4601）、Protocol Independent Multicast-Dense Mode协议无关多播-密集方式 PIM-DM（RFC3973）

-
- 虚拟专用网 VPN和网络地址转换 NAT

1- Virtual Private Network虚拟专用网VPN
   (1) IP地址紧缺，一个机构实际申请到的IP地址数往往远小于机构内主机数
   (2) 互联网很不安全，机构内并不是所有主机都要接入到外部互联网
   (3) 仅在机构内使用的计算机可由使用TCP/IP的机构自行分配IP地址

1. 本地地址和全球地址
   (1) 本地地址：仅在机构内使用的IP地址，可由机构自行分配
   (2) 全球地址：全球唯一的，必须向互联网管理机构申请的IP地址
   (3) 为避免二义性，RFC1918指出了一些private address专用地址只用作本地地址，互联网中所有路由器不对这些目的地址的数据报进行转发
2. 三类本地专用IP地址/reusable address可重用地址
   (1) 10.0.0.0~10.255.255.255，即10.0.0.0/8，称为A类，24 位块
   (2) 172.16.0.0~172.31.255.255，即172.16.0.0/12，称为B类，20位块
   (3) 192.168.0.0~192.168.255.255，即192.168.0.0/16，称为C类，16位块
   (4) 专用网/专用互联网/本地互联网：采用专用IP地址的互联网络
3. 虚拟专用网VPN：利用公用互联网作为本机构各专用网间的通信载体的专用网
   (1) 虚拟：并没有使用通信专线
   (2) 专用：仅供本机构内的通信

(3) 因而可能需要解决加密问题和分析实际IP的问题

4. 隧道实现VPN：进出外部互联网时封装为普通数据报，路由器使用全球地址
5. intranet内联网：机构内部网络所构成的VPN
6. extranet外联网：机构内和外部机构共同建立的VPN
7. remote access VPN远程接入VPN：由VPN软件建立与公司内部主机间的VPN隧道

2- Network Address Translation网络地址转换NAT

1. 是94年申请新IP地址已基本不可能时提出的，让专用网用户连接互联网的方法
2. 装有NAT软件的路由器称为NAT路由器，它至少要一个全球地址
   (1) 使用本地地址的主机与互联网主机通信时，要让NAT路由器代为转换地址，（并计入地址转换表）再连接互联网
   (2) 离开专用网时替换源地址为全球地址
   (3) 进入专用网时替换目的地址为内部地址
   (4) NAT路由器有n个全球地址时，专用网内最多只有n台主机能同时接入互联网，否则需轮流使用NAT路由器
   (5) 专用网内部的主机只能发起向外的通信，自然不能充当服务器
3. Network Address ans Port Translation网络地址与端口号转换NAPT
   (1) 将运输层的端口号也利用上，使更多主机能同时使用NAT路由器
   (2) 为区分，将老的NAT称作traditional NAT

-
- 多协议标记交换MPLS

1. IETF于1997年成立了 MPLS 工作组，开发出一种新的协议——多协议标记交换 MPLS (MultiProtocol Label Switching)
   (1) “多协议”表示在 MPLS 的上层可以采用多种协议，例如：IP，IPX；可以使用多种数据链路层协议，例如：PPP，以太网，ATM 等
   (2) “标记”是指每个分组被打上一个标记，根据该标记对分组进行转发
2. 为了实现交换，可以利用面向连接的概念，使每个分组携带一个叫做标记 (label) 的小整数。当分组到达交换机（即标记交换路由器）时，交换机读取分组的标记，并用标记值来检索分组转发表。 这样就比查找路由表来转发分组要快得多
3. MPLS 并没有取代 IP，而是作为一种 IP 增强技术，被广泛地应用在互联网中。
4. MPLS 具有以下三个方面的特点：
   (1) 支持面向连接的服务质量；
   (2) 支持流量工程，平衡网络负载；
   (3) 有效地支持虚拟专用网 VPN。

1- MultiProtocol Label Switching工作原理

1. IP 分组的转发
   (1) 在传统的 IP 网络中，分组每到达一个路由器后，都必须提取出其目的地址，按目的地址查找路由表，并按照“最长前缀匹配”的原则找到下一跳的 IP 地址（请注意，前缀的长度是不确定的）。
   (2) 当网络很大时，查找含有大量项目的路由表要花费很多的时间。

(3) 在出现突发性的通信量时，往往还会使缓存溢出，这就会引起分组丢失、传输时延增大和服务质量下降。

2. MPLS协议下的转发
   (1) 在 MPLS 域的入口处，给每一个 IP 数据报打上固定长度“标记”，然后对打上标记的 IP 数据报用硬件进行转发。
   (2) 采用硬件技术对打上标记的 IP 数据报进行转发就称为标记交换。
   (3) “交换”也表示在转发时不再上升到第三层查找转发表，而是根据标记在第二层（链路层）用硬件进行转发。
3. MPLS 域 (MPLS domain) 是指该域中有许多彼此相邻的路由器，并且所有的路由器都是支持 MPLS 技术的标记交换路由器 LSR (Label Switching Router)
4. LSR 同时具有标记交换和路由选择这两种功能，标记交换功能是为了快速转发，但在这之前LSR 需要使用路由选择功能构造转发表。
5. 转发等价类 FEC (Forwarding Equivalence Class)：路由器按照同样方式对待的分组的集合（“按照同样方式对待”表示：从同样接口转发到同样的下一跳地址，并且具有同样服务类别和同样丢弃优先级等）
   (1) 划分 FEC 的方法不受什么限制，这都由网络管理员来控制，因此非常灵活。
   (2) 入口结点并不是给每一个分组指派一个不同的标记，而是将属于同样 FEC 的分组都指派同样的标记。
   (3) FEC 和标记是一一对应的关系。
   (4) 负载平衡：网络管理员采用自定义的 FEC 就可以更好地管理网络的资源，这种均衡网络负载的做法也称为流量工程 TE (Traffic Engineering) 或通信量工程。

2- MPLS工作过程

1. MPLS 域中的各 LSR 使用专门的标记分配协议 LDP 交换报文，并找出标记交换路径 LSP。各 LSR 根据这些路径构造出分组转发表。
2. 分组进入到 MPLS 域时， MPLS 入口结点把分组打上标记，并按照转发表将分组转发给下一个 LSR。给 IP 数据报打标记的过程叫做分类 (classification)。
3. 一个标记仅仅在两个标记交换路由器 LSR 之间才有意义。分组每经过一个 LSR，LSR 就要做两件事：一是转发，二是更换新的标记，即把入标记更换成为出标记。这就叫做标记对换 (label swapping)。
4. 当分组离开 MPLS 域时，MPLS 出口结点把分组的标记去除。再以后就按照一般分组的转发方法进行转发。
5. 上述的这种“由入口 LSR 确定进入 MPLS 域以后的转发路径”称为显式路由选择 (explicit routing)，它和互联网中通常使用的“每一个路由器逐跳进行路由选择”有着很大的区别。

3- MPLS首部的位置与格式

1. 位置
   (1) MPLS 并不要求下层的网络都使用面向连接的技术。
   (2) 下层的网络并不提供打标记的手段，而 IPv4 数据报首部也没有多余的位置存放 MPLS 标记。

(3) 这就需要使用一种封装技术：在把 IP 数据报封装成以太网帧之前，先要插入一个 MPLS 首部。

(4) 从层次的角度看，MPLS 首部就处在第二层和第三层之间。

2. 首部格式

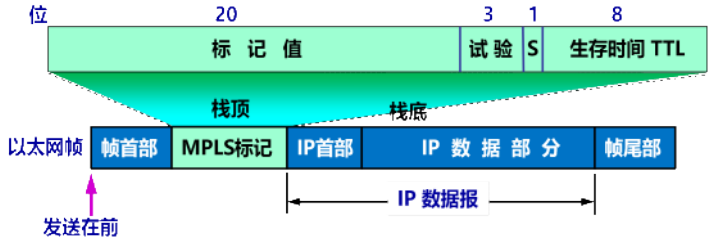
(1) 标记值（占 20 位）：可以同时容纳高达 220 个流（即 1048576 个流）。实际上几乎没有哪个 MPLS 实例会使用很大数目的流，因为通常需要管理员人工管理和设置每条交换路径。

(2) 试验（占 3 位）：目前保留用作试验。

(3) 栈S（占 1 位）：在有“标记栈”时使用。

(4) 生存时间TTL（占 8 位）：用来防止 MPLS 分组在 MPLS 域中兜圈子。

(5)



1)

2)

3)

4)

5)

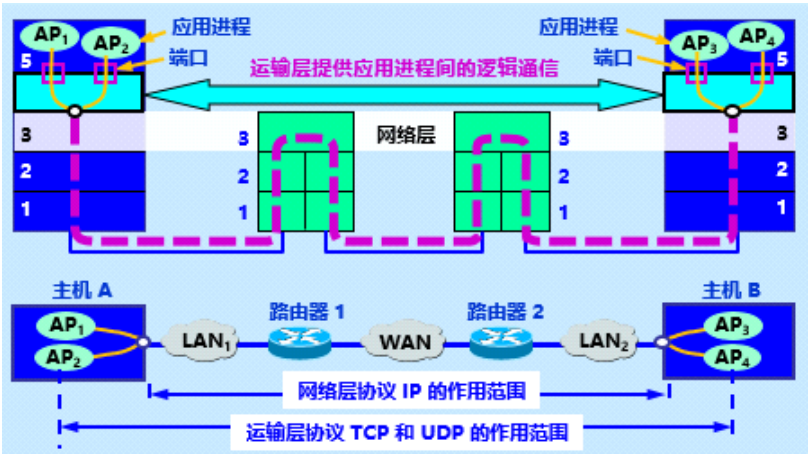
6) ----------我是底线----------

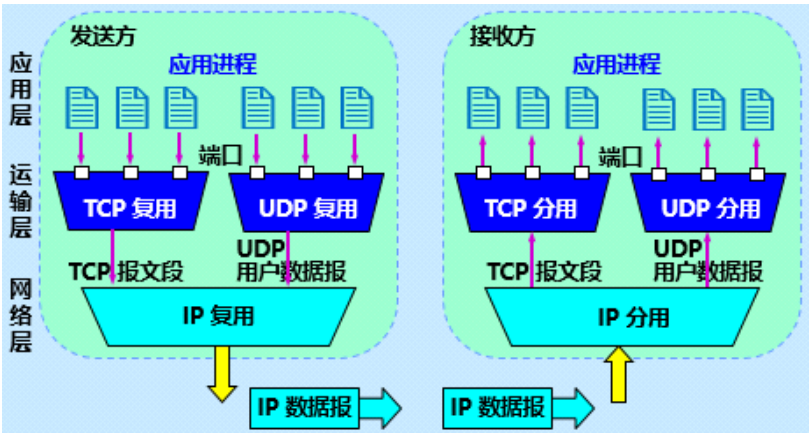
5运输层

2019年6月17日 1:49

-
- 运输层协议概述

1- 进程间的通信

1. 运输层的作用：面向通信部分的最高层，用户功能的最低层
   (1) 只有网络边缘终端主机才有运输层，转发只需前三层
   (2) 运输层视角的逻辑通信不是主机间的通信，而是进程间的
   (3) 

   (4) 运输层逻辑链路的复用和分用是基于端口的
   (5) 

   (6) 屏蔽作用：运输层向高层用户屏蔽了下层细节，仿佛端到端有信道

2- 两大运输协议

1. 无连接的User Datagram Protocol协议
   (1) 不可靠信道
   (2) 简单普适，支持单播多播广播
   (3) 用于DNS、DHCP、RIP、TFTP、SNMP、IGMP、NFS等
2. 面向连接的Transmission Control Protocol协议
   (1) 尽管下层网络一般是不可靠的，上层仍可视其为全双工可靠信道
   (2) 只支持单播
   (3) 用于HTTP、SMTP、FTP、TELNET等
3. Transport Protocol Data Unit运输协议数据单元TPDU：
   (1) TCP对应TCP segment报文段；UDP对应UDP报文或用户数据报

3- 运输层的端口

(1) 进程用进程标识符来标识，不同计算机的操作系统可能指派重复标识

(2) 进程的动态创建与撤销使双方难以互相识别

(3) 往往需要利用目的主机提供识别进程端点的功能，但不用知道实现方法

1. protocol port number协议端口号，简称端口

(1) 是逻辑意义上的软件端口，不同于路由器或交换机那种硬件端口

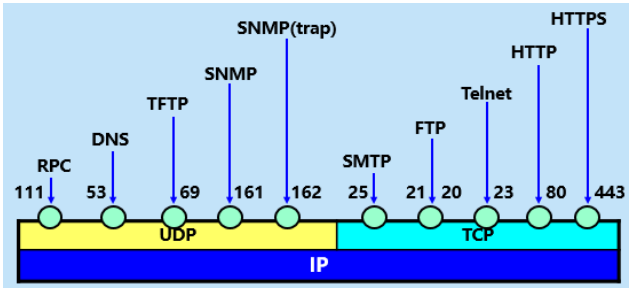
(2) 是16位的标志，允许65536个不同端口号

(3) 只有本地意义，标识计算机应用层中断各进程，不同计算机的相同端口号没有联系

2. 服务器的端口号

(1) well-known熟知端口：0~1023

1)



(2) 登记端口号：1024~49151

1) 为无熟知端口号的应用程序使用，必须在IANA登记，防重复

3. 客户端的端口号

(1) 又称短暂端口号：49152~65535

1) 留给客户进程暂时使用，通信完成后，重新向其他进程开放使用权

-
- 用户数据报协议UDP

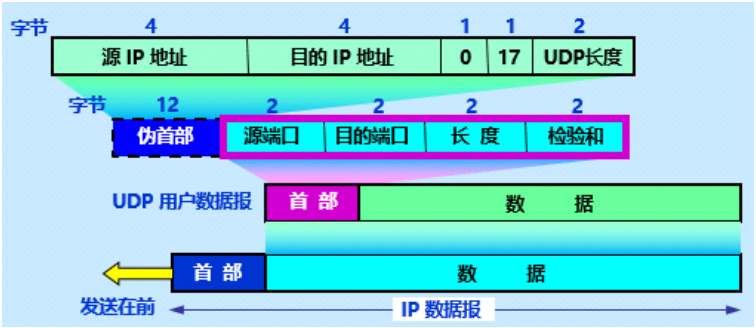
1- UDP概述

1. 只在IP数据报服务上增加了：复用分用功能和差错检测功能
2. 无连接，减少了建立连接的开销和发送前的时延
3. 尽最大努力交付，不可靠，不需要维持复杂的连接状态
4. 面向报文，不用拆分合并，保留报文边界，一次交付完（但可能在IP层被拆）
5. 没有拥塞控制，不会降低发送速率，适合多媒体通信
6. 支持一对一、一对多、多对一、多对多
7. 首部只有8字节，开销小（不过内容少时，IP层的首部占比会过大）

2- 首部格式

1. 由于UDP无连接，只需要用端口号，不需要套接字

(1)



2. 检验和
   (1) 计算检验和时会添加上伪首部，这一段是不会被传输到信道上的
   (2) 不同于IP数据报只检验首部，UDP数据报会带上数据一起算检验和
   (3) 以两字节为一行（空缺部分补零）求和得到的内容的反码填入检验和
   (4) 注意这个求和需要“回卷”，左边溢出的内容加回右边
   (5) 检验时，算上检验和，对UDP数据报求反码和，得到全1则视作无误

-
- 传输控制协议TCP概述

1- TCP主要特点
   1. 面向连接
      (1) 在无连接不可靠网络基础上提供可靠交付服务
      (2) 只能建立只有两个endpoint端点的一对一的全双工连接
      (3) 建立的是抽象的虚连接，并不是物理意义上的连接
   2. 面向字节stream流
      (1) 会把上层传下来的数据块视作无结构的字节流
      (2) TCP报文段不保证与程序数据块有对应关系
      (3) 会根据窗口值和网络拥塞程度来决定报文段应含几字节

2- TCP的连接
   1. socket套接字/插口
      (1) 端口号concatenated with拼接到IP地址即为套接字
      (2) 即socket=IP address:port
   2. 每条TCP连接由两个套接字确定
      (1) TCP连接::={socket1,socket2}={IP1:port1,IP2:port2}
      (2) 同一个IP地址可以建多个TCP连接，同一端口号也可出现在不同TCP连接
   3. socket同名的各种概念
      (1) 访问互联网的应用编程接口也叫socket Application Programming Interface，其中还有个函数也叫socket，调用该函数的端点也可称为socket，函数返回值也可称为socket描述符
      (2) 操作系统内核中联网协议的Berkeley实现也称为socket实现
      (3) RFC793定义的socket才是TCP连接用到的socket
         1)
         2)
         3)
         4)
         5)
         6)
         7)
         8) ----------我是底线----------

5可靠传输和首部

2019年6月18日 13:39

-
- 可靠传输的工作原理

1. 理想的传输条件：传输信道不产生差错；接收方总是来得及处理数据

1- 停止等待协议

(1) 早期数据链路层也会使用停止等待协议来保证尽量可靠

(2) TCP的等待方法复杂得多，以后的笔记以单向发送的情况做说明

1. 无差错情况

(1) 发送完一个分组（指TCP报文段）后暂停等接收方的确认消息

(2) 收到确认后才发送下一个分组

2. 出现差错

(1) 差错指接收方B检测到数据差错或未接收到丢失的分组

(2) 发送方A为发送的分组设置一超时计时器，若规定时间内未收到B的确认消息，则默认分组传送出现差错，重发该分组，直至收到确认

(3) 为防止网络延迟使B收到误以为丢失的重复发送的分组，A需给分组编号

(4) 为了能重传，发送后仍需保存副本，待收到确认后再清除

(5) 具体等待时间要考虑拥塞，时延，必须大于RTT，但又不适合太大

3. 确认丢失和确认迟到

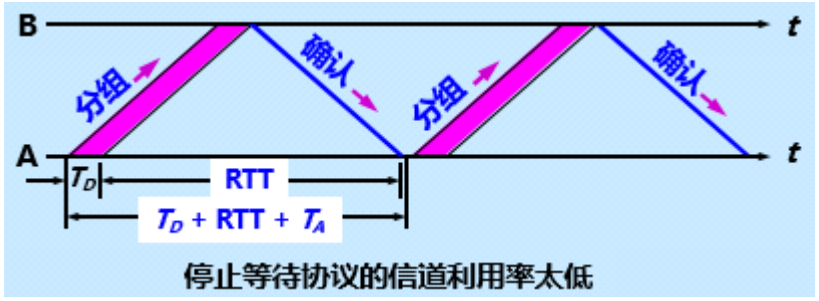
(1) B连续收到两个序号为M1的分组，首先要丢弃该分组，即不向上层交付

(2) 之后要重发确认，因为上一次确认可能丢失了，需要重新确认

(3) 若A连续收到两个对M1的确认，则说明是前一个说明迟到了，需忽略后一个；若反复收不到确认，说明通信线路实在太差

(4) 因而该协议又称为Automatic Repeat reQuest自动重传请求ARQ，因为任何情况下接收方都不会主动请求重传某一分组，全靠发送方自行猜测

4. 信道利用率

(1) 

(2) 信道利用率U=TD/TD+RTT+TA

(3) 其中TD=发送分组时间=数据长度/发送速率；TA=发送确认时间；RTT=分组传送时间+确认传送时间（严格来讲分子应该不包含首部的时间，但一般利用率不考虑那么严格）

(4) 当RTT远大于TD及反复重传时，信道利用率就极低

(5) 为了提高利用率，显然应采用流水线传输，使用连续ARQ和滑动窗口

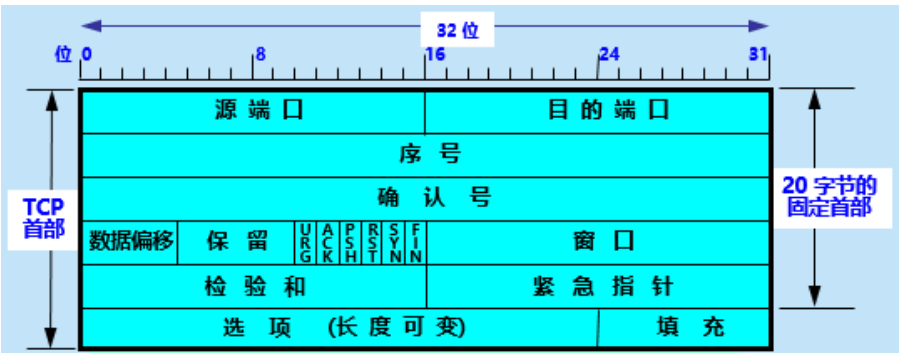
2- 连续ARQ协议

1. 滑动窗口思想：将可发送的首个和末个分组视作窗口首尾
   （1）每收到一个确认，A就把发送窗口向前滑动一下（序号大的方向）
2. B的确认方式一般是累积确认：按序到达的最后一个分组号
3. A的重传方式一般是go-back回退：从第一个B未收到的分组号开始重传
4.

	连续ARQ协议	停止等待协议
发送的分组数量	一次发送多个分组	一次发送一个分组
传输控制	滑动窗口协议	停等-等待
确认	单独确认 + 累积确认	单独确认
超时定时器	每个发送的分组	每个发送的分组
编号	每个发送的分组	每个发送的分组
重传	回退N，多个分组	一个分组

-
- TCP报文段的首部格式

1- 概述

1. 

2. 前20字节是固定的（最小长度），后4n字节是可选项

2- 固定部分

1. 源、目的端口；各2字节
2. 序号seq；4字节；~2^32 - 1；共2^32=4G=4294967296个
3. 确认号ack；占4字节；期待对方下一个报文段的序号；若确认号=N时，表示0~N-1字节的数据都正确收到了
4. 数据偏移；占4位；5~15；数据段的起始位置，即首部长度；单位是32位，说明TCP首部最大长度40字节
5. 保留；占6位；暂时没想好干啥，都置0
6. 紧急URGent；1位；取1时表示有紧急数据，如中断命令，需要插队
7. 确认**ACK**nowledgment；1位；取1时表示是确认报文，确认号才有效
8. 推送PuSH；1位；取1时表示希望对方不等缓存填满即直接向上层交付
9. 复位ReSeT；1位；取1时表示出现严重差错，希望对方释放TCP连接
10. 同步**SYN**chronization；1位；取1而ACK取0时表示这是这是连接请求，取1且ACK也取1时表示是接受连接报文
11. 终止**FIN**ish；1位；取1时表示数据发送完毕，请求释放连接
12. 窗口rwnd；占2字节；0~2^16 -1；表示该报文的发送方的接受窗口大小
13. 检验和；占2字节；类似UDP，加上12字节的伪首部（协议字段不同于UDP的

17，应该是6，如果是IPv6，伪首部也会对应改）

14. 紧急指针；占2字节；紧急位取1时，用于指出紧急数据的字节数，即末尾在报文段中的位置；神秘的是窗口为0时照样可以发紧急数据
15. 选项；长度0~40字节；没有选项的首部长度，即前14项的长度=20字节
16. 填充字段；当首都长度不为4字节整数倍时填充的0

3- Maximum Segment Size最大报文段长度：数据字段的最大长度

1. MSS+首部长度=真实TCP报文段最大长度
2. MSS默认值为536字节（没算上TCP和IP各至少20字节首部）

4- 选项

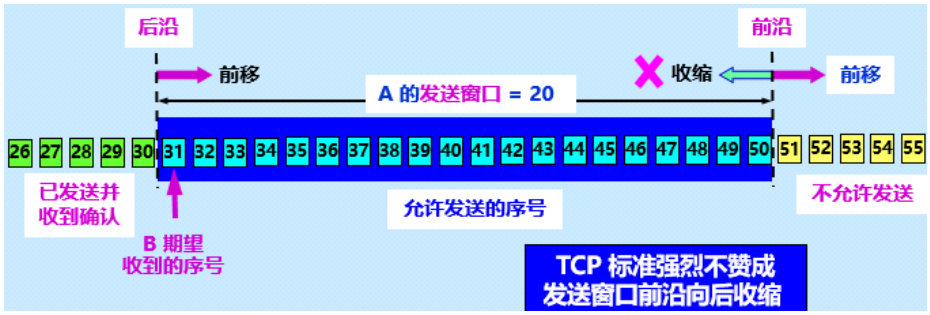
1. 窗口扩大；占3字节；有1字节表示移位值S，表示希望对方窗口大小的位数变成16+S，即0~2^(16+S)-1，S最大值为14
2. 时间戳；10字节；取其中有4字节时间戳值字段和4字节时间戳回送回答字段，用于计算RTT，和Protect Against Wrapped Sequence numbers，即防止序号绕回PAWS：避免序号seq超过2^32时无法判断序号有没有绕回
3. 选择确认SACK，详见下节最后

-
- TCP可靠传输的实现

1- 以字节为单位的滑动窗口

1. 发送窗口：未收到确认的情况下，可以发送的连续字节的数据的范围

(1)

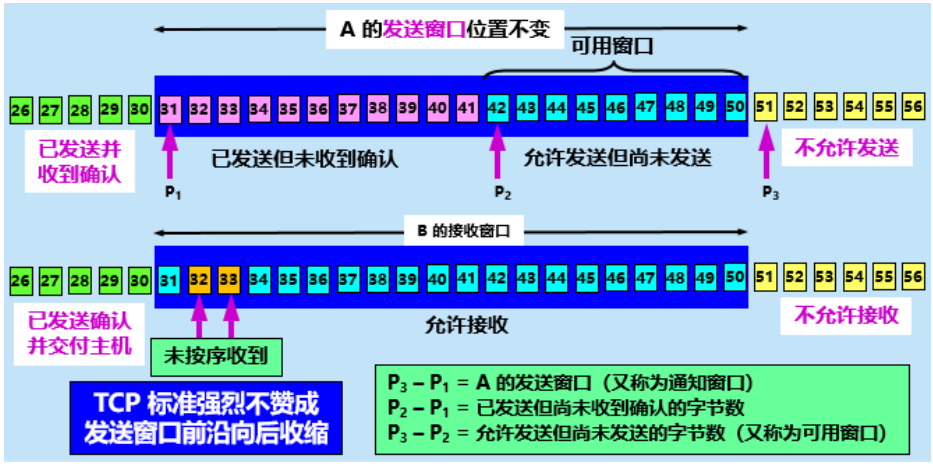


(2) 通常是发送缓存的子集

(3) 大小随对方的接收窗口大小和拥塞情况适当调整

2. 接收窗口：允许接收的数据的范围

(1)



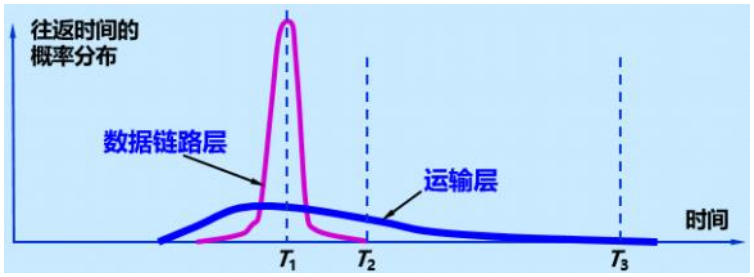
(2) 不按序到达的数据一般也不能交付给上层应用

(3) 连续到达的确认号可以稍晚发送，也可以在给对方传送数据的报文里捎带上，但最好不要晚超过0.5秒，避免不必要的重传

3. 发送缓存
    (1) 暂存打算发的数据
    (2) 暂存发完尚未收到确认的数据
    (3) 发送缓存的后沿和发送窗口的后沿是重合的
4. 接收缓存
    (1) 暂存尚未被程序读取的数据
    (2) 暂存未按序到达的数据

2- 超时重传时间的选择

1. 往返时间RTT

(1) 

2. 平滑的加权平均往返时间RTTs（s指smoothed）
    (1) 新RTTs=（1-α）|旧RTTs + α 新RTT
    (2) α越接近1会使RTT更新越快，RFC6298推荐α=1/8
3. RTTD=RTT偏差的加权平均值，与RTTs和新RTT样本的差的绝对值有关
    (1) 新RTTD=（1-B）|旧RTTD + β | RTTs-新RTT |
    (2) β的推荐值=1/4
4. 自适应算法
    (1) RetransmissionTime-Out超时重传时间RTO，在RFC6298中被推荐为
    (2) RTO=RTTs + 4 x RTTD
5. Karn算法
    (1) Q：重传报文的确认应该按首次发送还是重传发送计算RTT
    (2) A：重传了就不作为RTT样本
    (3) Q：反复重传导致无法更新RTO，再导致更多反复重传怎么办
    (4) A：每重传一次就让RTO乘以γ一次，γ的典型值是2

3- Selective ACK选择确认SACK

1. Q：两段较长连续数据间缺了一个序号的数据，能不能不让后段数据被重传
2. A：可以在首部选项里添加SACK，说明[L,R)才是需要重传的内容
3. 需要在建立连接前讲清楚要不要使用这个功能
4. RFC2018对[L,R)边界格式有详细的规定，但并没有要求对方该怎么处理SACK
    1)
    2)
    3)
    4)
    5)
    6) ----------我是底线----------

5流量控制和连接管理

2019年6月18日 20:39

-
- TCP的流量控制

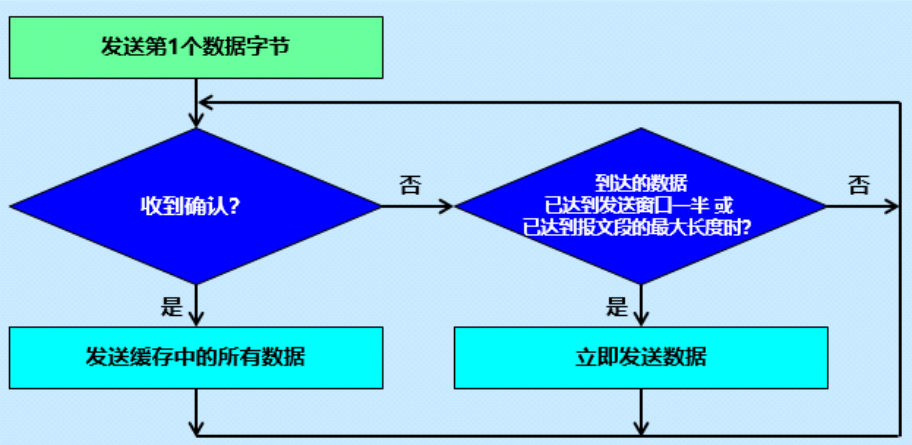
1- 滑动窗口实现流量控制

1. flow control流量控制：让发送速率不要太快，让接收方来得及接收
   (1) 注意窗口单位是字节，不是报文段序号
   (2) ACK表示确认位，ack表示确认字段值，rwnd表示接收窗口大小
2. 接收方B接收缓存不够了以后，即可向发送方A更新rwnd=0的报文
   (1) 这个报文称为零窗口通知
   (2) A收到后也会立刻停止继续发送
   (3) 不过发送零窗口通知后也有需要接收的临时报文：零窗口探测报文，确认报文，紧急数据报文段
3. B上交了一定量的数据后，即可向A更新rwnd为新的较大值的报文
   (1) 若之前给了A零窗口通知，现在这个更新rwnd的报文却丢失了，则会进入死锁局面，A不敢发，B收不到
   (2) 为解决死锁的可能，TCP为每个连接设置一个persistence timer持续计时器，定时“探测”：发送一个仅携带1字节数据的零窗口探测报文段，要求对方更新rwnd值

2- TCP的传输效率

1. 发送报文的时机
   (1) 当缓存内字节数达到MSS时就组装成一个报文段并发送
   (2) 当对面发来PSH请求时，直接发送
   (3) 发送计时器的期限到了，将缓存内数据装入报文段发送
2. Nagle算法
   (1) silly window syndrome糊涂窗口综合症：RFC813：接收窗口字节数很少时，信道利用率和有效数据传输效率变得极低
   (2) Nagle算法：发送方第一次只送一个字节，缓存后序的所有数据，当收到确认报文后再将缓存全发送出去；之后也是，收到前一个报文的确认后才发下一个报文段；接收方的接收缓存到达swnd的一半或已到达MSS时也可以直接发确认报文，通知对方可以发送下一个报文段

(3)

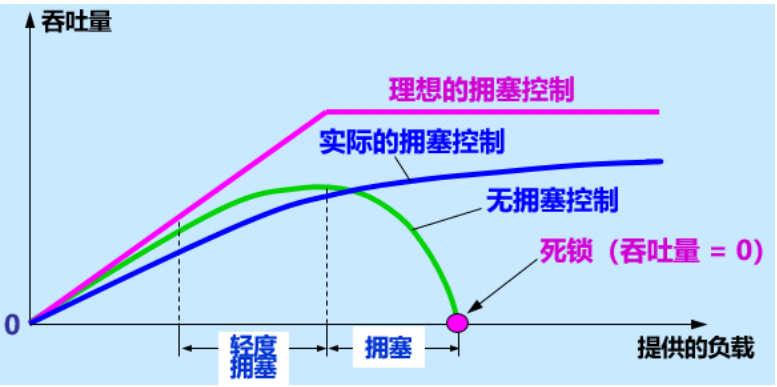


-
- TCP的拥塞控制

1- 拥塞控制的一般原理

1. congestion拥塞：总要求资源>可用资源
2. 产生原因：缓存容量小；链路容量小；处理机处理速率小；拥塞本身加剧拥塞
3. 拥塞控制：防止过多数据注入网络，使链路和路由器不过载；是全局管理
   (1) 流量控制：防止接收端来不及接收；是端到端的通信量控制问题

(2)



   (3) 上图横轴offered load又称输入负载或网络负载，表示单位时间内输入进网络的分组数目
   (4) 纵轴是throughput吞吐量，代表单位时间内从网络输出的分组数目
   (5) 拥塞是动态的问题，拥塞控制本身甚至能引起网络性能恶化甚至死锁
4. 解决方法
   (1) Q：增加资源能解决吗
   (2) A：不能。盲目增大缓存而不增处理速度可能使排队等待时间增加，引起大量超时重传；只提高处理机速率会将瓶颈转移去他处
   (3) **开环控制**：设计网络时力争考虑周全，不发生拥塞
   (4) 闭环控制：基于**反馈环路**，按网络运行状态采取相应控制措施
5. 闭环控制措施
   (1) 监测网络系统，检测拥塞在何时何处发生
      1) 主要监测：缺缓存而丢弃分组的比例；平均队列长；超时重传分组数；平均分组时延；分组时延的标准差
   (2) 将拥塞发生的信息传送到可采取行动的地方
      1) 传递拥塞通知：通知拥塞发生的分组；在分组中保留表示拥塞状态的字段；周期性地发出探测分组

(3) 调整网络系统的运行以解决出现的问题

1) 调整过于频繁会使系统产生不稳定的振荡；迟缓采取行动又无价值

2) 调整思路：增加网络可用资源；减少用户对资源的需求

2- TCP的拥塞控制方法RFC5681

1. congestion window拥塞窗口cwnd

(1) 慢开始和拥塞避免是两个基于cwnd的拥塞控制

(2) 发送方让自己的**发送窗口=拥塞窗口**

1) 严格地说，发送窗口上限=min{对方的接收窗口，自己的拥塞窗口}

2) 根据网络拥塞程度，动态改变拥塞窗口大小

3) 按现在通信线路的质量，因传输差错而丢弃分组的概率<1%

4) **判断拥塞的依据是出现超时**，超时指一直收不到某序号的确认

5) 判断出拥塞后，一般立即设置拥塞窗口**cwnd=SMSS**

2. slow-start慢开始：cwnd值从小逐渐增大，试探网络负载能力

(1) cwnd每次增量=min（N，SMSS）其中N是原先未被确认的、但现在被刚收到的确认报文段所确认的字节数（应该是与首部ack有关）；SMSS是Sender Maximum Segment Size发送方最大报文段

(2) cwnd初值不推荐超过2~4个SMSS或超过6570字节

(3) cwnd/SMSS=每一transmission round传输轮次可发送的报文数

(4) 实际应用上，每次收到了B的确认报文，A就可多发1个报文，即让cwnd增加1SMSS

(5) 一般来说，每经过一传输轮次，cwnd都会翻倍

(6) 所以cwnd增速极快，“慢”开始并不是指增速慢，其实是指刚开始发送的报文很小，不会突然给网络增加很多负担

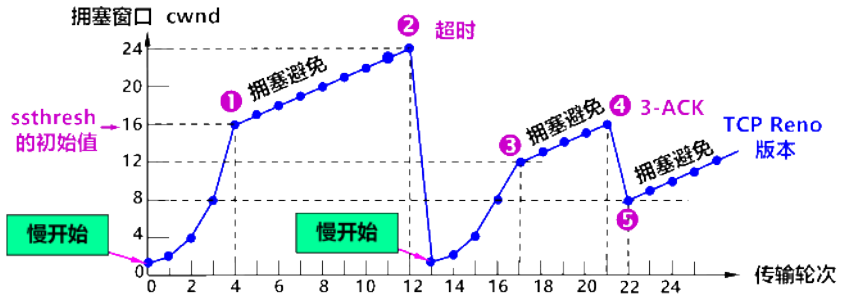
3. congestion avoidance拥塞避免：cwnd值线性缓慢增长

(1) 每次收到确认报文，就让cwnd增加1SMSS*SMSS/cwnd，即多发SMSS/cwnd个报文，则一个传输轮次下来，cwnd恰增加了1SMSS

(2) additive increase**加法增大**AI，拥塞避免依旧会增加拥塞的概率，此处“避免”只是指长得慢，不易出现拥塞

(3) 下图纵轴单位为SMSS，ssthresh初值为16个SMSS，Reno是指新版本，区分Tahao版本

(4)



当 TCP 连接进行初始化时，将拥塞窗口置为 1。图中的窗口单位不使用字节而使用报文段。

4. ssthresh慢开始门限：确认是否用慢开始算法更改cwnd的门限

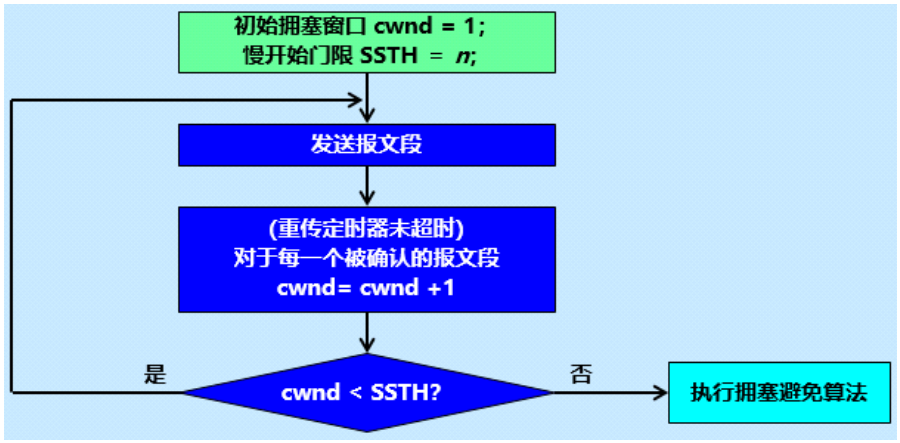
(1) cwnd<ssthresh时用慢开始算法，快速增长cwnd

(2) cwnd>ssthresh时用拥塞避免算法，缓慢增长cwnd

(3) cwnd=ssthresh时两个都能用

(4) 发现超时未确认时，令**ssthresh=max（cwnd/2，2）**

(5)



5. fast retransmit快重传：B发现个别报文段丢失后，立即向A发送确认报文

(1) A接到该确认后，会立即进行重传，实现快重传

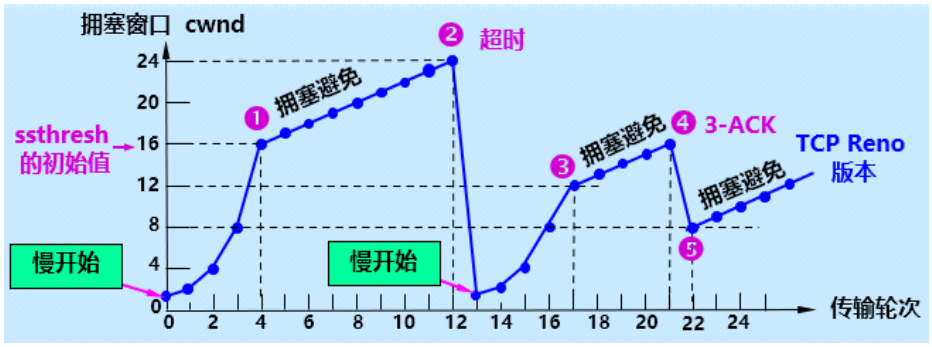
(2) A连续收到3个对同一序号的ACK，说明该序号后面的一个序号丢失或迟到的情况下，更后面的连续三个序号都到达了，因此需要快速重传

6. fast recovery快恢复：令**cwnd=ssthresh=cwnd/2**

(1) 用于检测到重复三连ACK时，说明需要快重传，但不一定是出现了拥塞，因此降低一半cwnd，再用加法增大试探

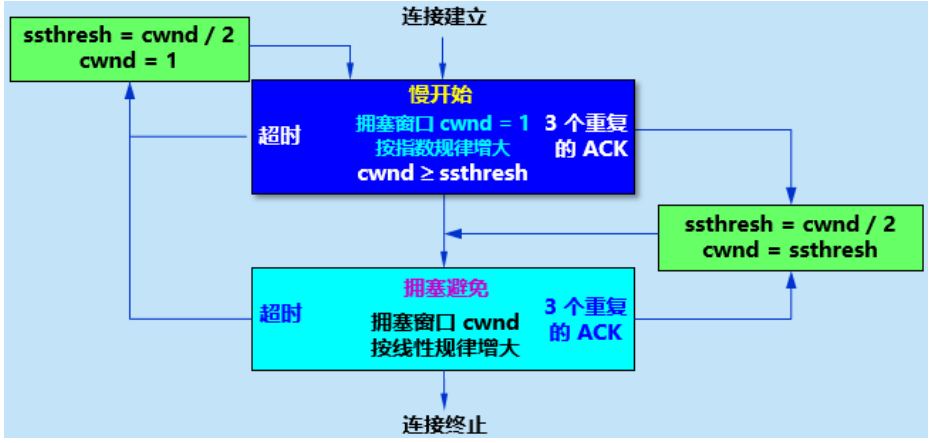
(2) 这个除以2被称作multiplicative decrease**乘法减小**MD

(3)



(4) AI和MD统称为AIMD算法

(5)



3- Active Queue Management主动队列管理AQM

1. 传统的TCP报文丢弃策略：先进先出FIFO，tail-drop policy尾部丢弃

(1) 丢弃路由器队列尾部往往会导致一连串分组的丢失，迫使发送方超时重

传，再慢开始

(2) global syncronization全局同步：许多TCP连接同时进入慢开始状态，使全网通信量急降，网络恢复正常后再突增

(3) 98年提出来主动队列管理，当队列长度到达某一值得警惕的数值时，主动丢弃分组

2. Random Early Detection随机早期检测RED（D还可以是Drop或Discard）

(1) 当网络中平均队列长度<最小门限，则直接排队；若>最大门限，则直接丢弃；若处于之间，则按概率p丢弃

(2) 2015年的RFC7567已不推荐RED，但AQM还是需要的

-
- TCP的运输连接管理

1. TCP连接三个阶段：连接建立、数据传送、连接释放

1- TCP的连接建立

1. 建立连接的三个问题

(1) 双方互相确认对方存在

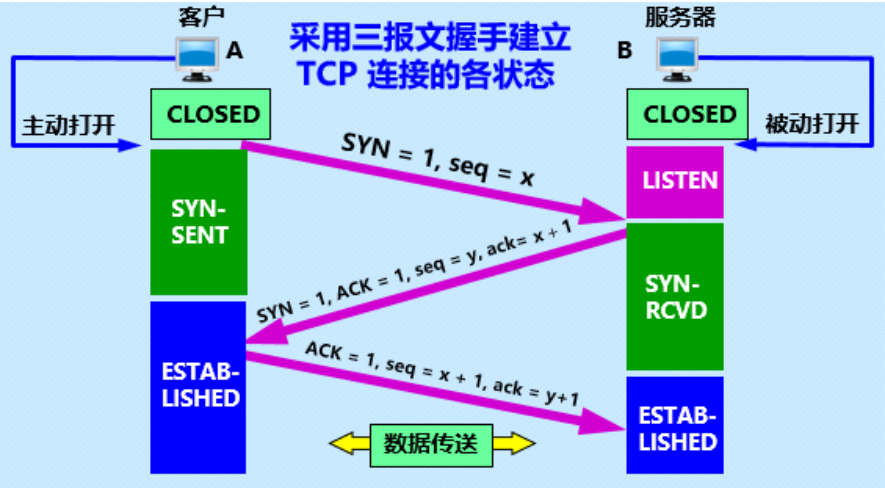
(2) 协商参数（窗口最大值，是否使用窗口扩大、时间戳选项、服务质量）

(3) 对运输实体资源进行分配（缓存大小，连接表中的项目等）

2. 连接建立：三报文握手

(1) TCP连接的建立采用客户服务器方式，主动发起连接的A视作client客户，被动等待连接建立的B视作server服务器

(2) 三报文主要是为了**防止迟到的连接请求引起不必要的连接建立**

(3) 

(4) 客户A先主动发出请求（SYN=1），并进入SYN-SENT状态

(5) LISTEN状态的B的服务器进程收到该报文后，创建一个Transmission Control Block传输控制块TCB，再发送同样SYN=1的第二条报文，进入SYN-RCVD状态

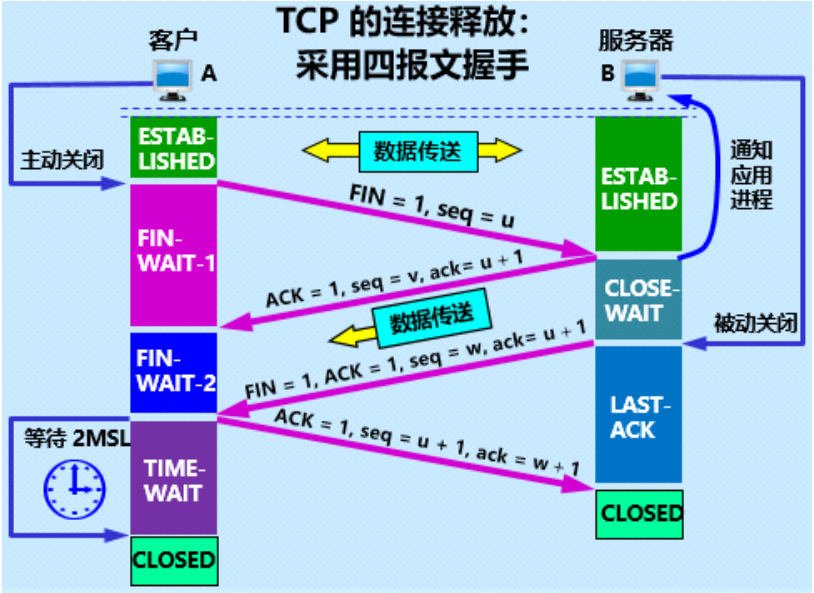
(6) A在收到这第二个报文后进入ESTABLISHED，并立刻发送第三条报文（ACK=1）；B通过第三条报文确认第二条报文被A收到后，也正式ESTABLISHED

(7) 第二条报文同时负责了确认收到第一条报文（ACK=1，ack=x+1）和请求对方确认无误（SYN=1，seq=y）两条功能，可以拆成两次发，成为四报文握手

2- TCP的连接释放

1. TCP规定释放连接请求报文段的FIN=1，且该报文必须消耗一个序号

(1)



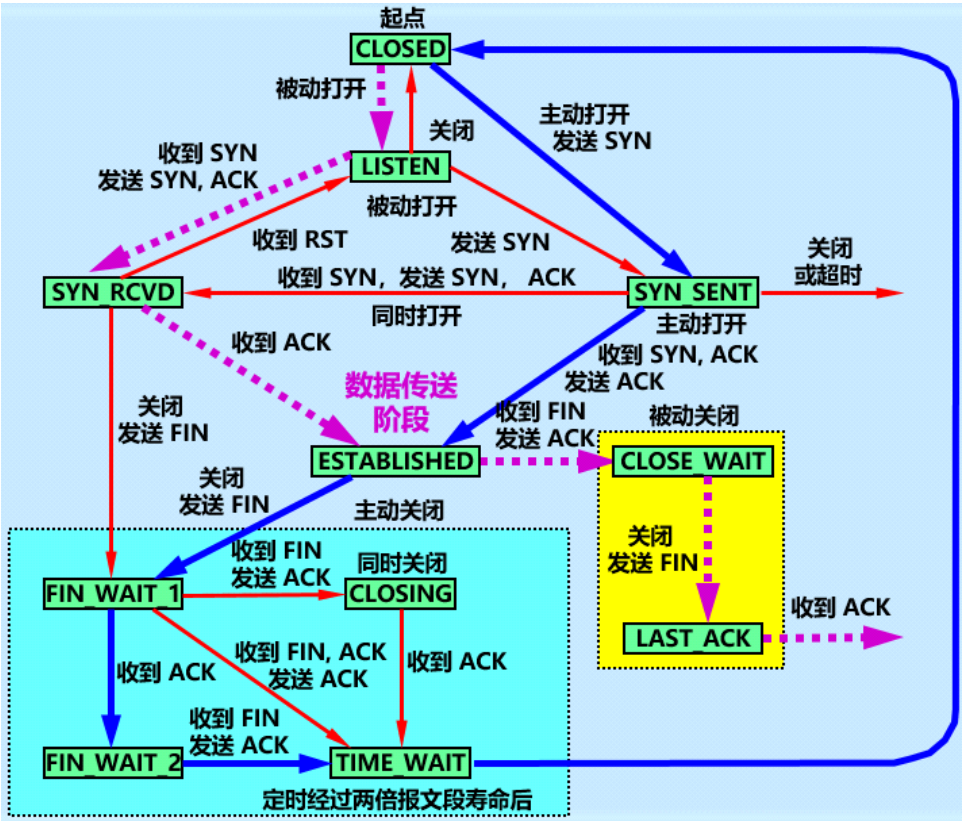
2. 释放连接的四报文
   (1) A先发送释放请求（FIN=1，seq=u，u-1为A之前发的最后一个数据字节的序号）并进入FIN-WAIT-1状态，等待B的确认
   (2) B收到第一条报文后，先给A发送确认（ACK=1，seq=v，ack=u+1，v-1为B之前发的最后一个数据字节的序号）并进入CLOSE-WAIT状态，并通知上级应用进程
   (3) A收到B发来的请求后进入FIN-WAIT-2状态，等待B也打算关闭连接
   (4) B的进程确认没有想发给A的数据时，发送释放请求（FIN=1，ACK=1，seq=w，ack=u+1，w-1是B之前发的最后一个数据字节的序号，u是A的释放请求的序号）并进入LAST-ACK状态，等待A确认关闭
   (5) A收到后发出确认（ACK=1，seq=u+1，ack=w+1）进入TIME-WAIT状态，等待2Maximum Segment Lifetime时间后正式CLOSED，并撤销相应的TCB
   (6) B也在收到该请求后撤销TCB，释放连接
3. 等待2倍MSL时间的原因
   (1) **保证第四条报文段不丢失。**这是通过B迟迟收不到确认时会发重传第三条报文来实现的，A收到这个重传时，重新发第四条报文并重启计时
   (2) 确保网络中没有迟到的报文，如失效的连接请求报文，为防止它影响下次连接，多等一会
4. keepalive timer保活计时器
   (1) 客户端主机突然故障，没必要继续保持连接，但又发不出释放连接请求
   (2) 需要服务器自动判断有没有必要主动断开连接
   (3) B每收到A一次报文，就重启一次计时器，若计时器达到0，就每隔75秒发送一探测报文段，连续发送10个后若仍无响应，则主动断开连接

3- TCP的有限状态机

   (1) 粗实线表示客户进程正常变迁
   (2) 粗虚线表示服务进程正常变迁

(3) 细线表示异常变迁

1.



(1) 细线异常可能和分组丢失有关，分组丢失除了在拥塞状况下会发生外，还在这些情况下发生：

1) IP层某分片迟到，终点来不及组装，全部丢失
2) 到达终点，终点缓存不够了
3) 路过某局域网的网桥，网桥空间不够了
4)
5)
6)
7) ----------我是底线----------

6DNS，FTP和TELNET

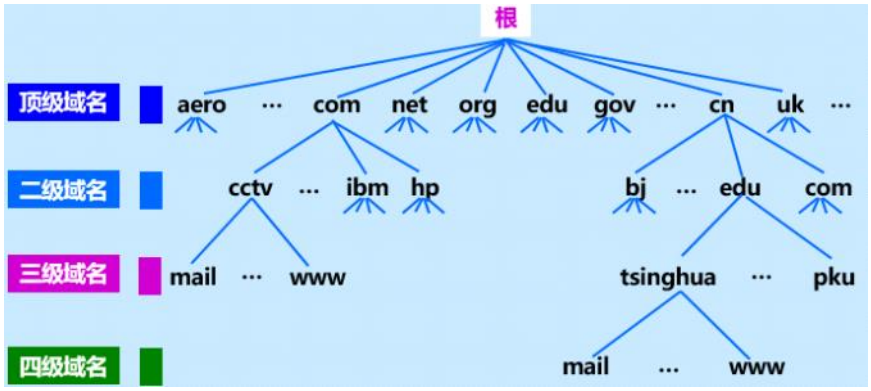
2019年6月19日 14:43

-
- 域名系统DNS

1- Domain Name System概述

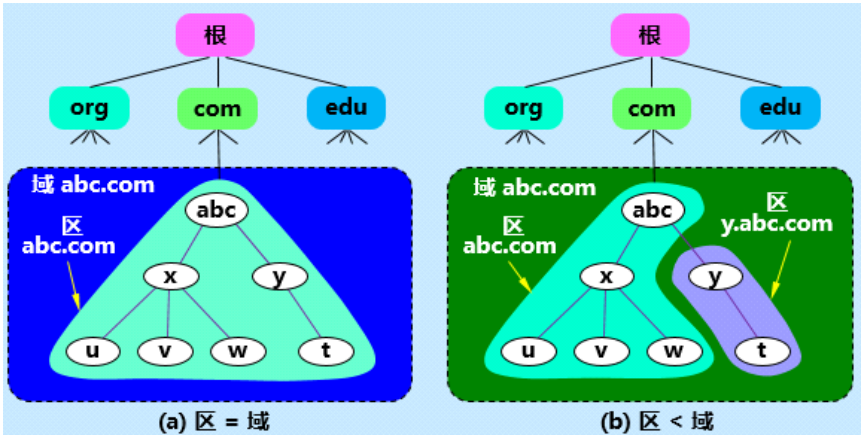
1. 互联网的命名系统中使用了许多domain域，域名系统是用于互联网中，让计算机用户把主机名转换为IP地址。RFC1034，1035
2. DNS大多数名字都是本地resolve解析，因为DNS是分布式系统，部分主机的故障不会妨碍整个系统的正常运行
3. 需要解析主机名的应用进程会调用resolver解析程序，并成为DNS的一个客户，把待解析域名填入DNS报文，用UDP数据报发给本地域名服务器
4. 本地域名服务器若回答不了，就自己也成为一个客户，询问其他域名服务器

2- 互联网的域名结构

1. 是一种层次结构
   (1) 每个域名都是一个label标号组成，不区分大小写英语字母
   (2) 标号间用英语.构成
   (3) 域名的级别从左到右递增
   (4) 整个域名，不能超过255个字符
   (5) 注意这个逻辑概念上的层次和物理意义上的存放位置无关
   (6) 

2. 三类top level domain顶级域名
   (1) 国家顶级域名nTLD，.cn 表中国，.us 表美国，.uk 表英国等
   (2) 通用顶级域名gTLD：.com表公司企业.net表网络服务机构.org表非盈利性组织.edu表美国教育机构.gov表美国政府.mil表美国军事.int表国际组织等7个，后又增加了13个.aero表航空运输企业.biz表公司和企业.cat表加泰隆人的语言和文化团体.coop表合作团体.info表各种情况.jobs表人力资源管理者.mobi表移动产品与服务的用户和提供者.museum表博物馆.name表个人.pro表有证书的专业人员.travel表旅游业
   (3) infrastructure domain基础域名：arpa，用于反向域名解析
   (4) 2011年ICANN又开放了新顶级域名申请，中文顶级域名都出现了60个
3. 域名服务器
   (1) 一个服务器管辖一个zone区，称为其authoritative name server权限域

名服务器，区一般是域的子集

1) 

2) 一般数据除了存在master主服务器外，还会定期把副本存在secondary辅助服务器

(2) 四大类域名服务器

1) root name server根域名服务器：知道所有顶级域名服务器的域名及IP地址，是最重要的服务器，16年已有约600个根服务器，共享有13个不同IP地址的域名a~m.rootserver.net，即anycast任播技术：不同主机使用同一IP，同时作为通信终点。因此DNS客户只需就近选择一个根域名服务器即可解析。RFC2870

2) TLD顶级域名服务器：管理对应顶级域名的所有二级域名

3) 权限域名服务器：之前提的负责一个区的域名服务器，当给不出回答时，会返回下一个应查找的域名服务器

4) local本地域名服务器（默认域名服务器）：并不在域名服务器层次结构，一般只是作为代理，成为DNS客户，去查其他域名服务器，仅当请求解析主机位于同一ISP时能直接返回转换结果

(3) 高速缓存：为减轻根域名服务器负担，其他服务器最好都在本地缓存最近一段时间（如两天）的解析结果，以便减少对其他服务器的询问

(4) 服务器高速缓存内无转换结果时的解析过程

1) recursive query递归查询：让域名服务器以DNS客户身份向其他域名服务器发出解析请求（较少用）

2) iterative query迭代查询：返回下一个该查询的域名服务器，让客户去向它提出解析请求

-
- 文件传输协议FTP

3- File Transfer Protocol概述

1. 是互联网最被广泛使用的复制整个文件的传送协议，RFC959
2. 交互式的访问，允许客户指明文件类型与格式，允许存取权限
3. 屏蔽了不同计算机系统的细节，透明存取
4. online access联机访问：多个程序可同时对同一文件进行存取
5. 属于文件共享协议的Network File System正在建议标准里待着，远程操控文件，暂时只能用FTP整个读过来，再整个拷回去

4- FTP基本工作原理

1. 需要解决的问题
   (1) 存储数据格式不同
   (2) 目录结构和命名规定不同
   (3) 存取功能对应的操作系统命令不同
   (4) 访问控制不同
2. FTP服务器
   (1) 主进程：打开熟知端口21，等待到客户请求时，建立从进程
   (2) 从进程：处理单个请求，分为控制进程和数据传送进程
3. 客户和服务器间的连接
   (1) 控制连接：会话期间保持打开，如接受终止请求，与传送链接并行
   (2) 传送连接：用于传输文件的连接，端口号是20
   (3) 这种分离控制称为控制信息out of band带外传送

5- Trivial FTP简单文件传送协议

1. 也是客户服务器方式，但使用UDP数据报，端口号69
2. 优点
   (1) 不需要建立连接
   (2) 不占内存
3. 缺点
   (1) 只支持文件传输，不支持交互
   (2) 无庞大命令集，无列目录功能，无身份鉴别
4. 特点
   (1) 除最后一次外，固定传512字节，非512的视作结尾（若结尾恰512，则再发一个空的）
   (2) 按报文序号编号，1开始
   (3) 支持ASCII码和二进制传送
   (4) 可对文件读写
   (5) 首部简单
   (6) 类似停止等待协议

-
- 远程终端协议TELNET

1- TELNET又称终端仿真协议

1. 换行符在某些系统是ASCII的CR回车，有些是LF换行，有些是CR和LF；中断在某些系统是ctrl+c，在某些是ESC
2. 为解决不同系统的不同命令传输，定义了Network Virtual Terminal网络虚拟终端NVT
3. 服务器负责监听远低系统传来的控制信号，把格式转换为NVT格式；本地客户负责转换回本地所需格式
4. 协议双方是平等的，可以有option negotiation选项协商

   1)
   2)

3)
4)
5)
6)
7)
8) ----------我是底线----------

6WWW

2019年6月19日 16:26

-
- 万维网WWW

1- World Wide Web概述

1. 是大规模的、联机式的信息储藏所，可简称为Web
   (1) 工作方式是客户-服务器
   (2) 服务器运行服务器程序
   (3) 客户端运行浏览器等客户程序
2. 是一个分布式的hypermedia超媒体系统，是hypertext超文本系统的扩充
   (1) 超文本是指由link指向其他文档的文档，是万维网的基础
   (2) 采用链接的方法，使万维网能轻松地从一个站点访问到其他站点
   (3) 而hypermedia超媒体文档除了文本外还包含图形、图像、声音、动画、视频等，算是超文本系统的扩充
   (4) 客户程序主窗口上显示的万维网文档称为page页面
3. 需要解决的问题
   (1) Q：如何标志分布在整个互联网上的万维网文档
   (2) A：Uniform Resource Locator统一资源定位符URL
   (3) Q：如何实现各种链接
   (4) A：HyperText Transfer Protocol超文本传输层协议HTTP配合TCP
   (5) Q：如何显示各种万维网文档
   (6) A：HyperText Markup Language超文本标记语言HTML
   (7) Q：如何方便查找所需信息
   (8) A：搜索引擎

2- 统一资源定位符URL

1. URL的格式：<协议>**://**<主机>:<端口>/<路径>
   (1) 协议和主机都不区分大小写
   (2) 路径可能随操作系统不同而区分大小写
   (3) 常见协议：ftp，http，news，对多数浏览器来说，可省
   (4) 主机是指域名，最前的最小域名ww一般可省，后面的域名就没法省了
   (5) 端口号和路径都可省
2. 使用HTTP的URL
   (1) 格式：http://<主机>:<端口>/<路径>
   (2) 端口号是80，一般可省
   (3) 路径项也缺省时，会访问home page主页

3- 超文本传送协议HTTP

1. HTTP的操作过程
   (1) HTTP是transaction-oriented面向事务的，即有原子性，不能只传一部分信息，必须可靠地交付全部超媒体文档

(2) 服务器进程不断监听端口号80，建立TCP连接后，返回客户端每个请求对应的页面

(3) 请求和响应的交互格式即为HTTP的内容

1) 每次交互都由用户端的一个ASCII码串构成的请求和类似通用互联网扩充（MIME-like）的响应构成

2) 虽然使用了TCP连接来保证数据可靠，但HTTP协议本身无连接

3) HTTP1.0是stateless无状态的，同一客户每次访问统一服务器的文档都被视作新用户对待

(4) persistent connection持续连接

1) HTTP1.1为解决1.0每次请求都重新连接的浪费，默认使用持续连接，一段时间内不撤销连接，这个设置可以在浏览器的Internet选项中取消

2) without pipelining非流水线方式：收到前一个响应后才能发出下一个请求，每请求一个对象之后都要等一个RTT

3) with pipelining流水线方式：用户连续发送一串请求，服务器也连续发送一串响应，减少了空闲时间，提高了下载效率

2. proxy server代理服务器/Web cache万维网高速缓存

(1) 浏览器请求互联网服务器时，先和代理建TCP连接，若代理有最近申请过该文档，则直接从缓存里找出来，发给客户

(2) 缓存里没有该对象的话，才去跟origin server源点服务器建TCP连接，将目标放入缓存，再返给客户

(3) 代理服务器既充当服务器，又当客户（向源点申请资源时）

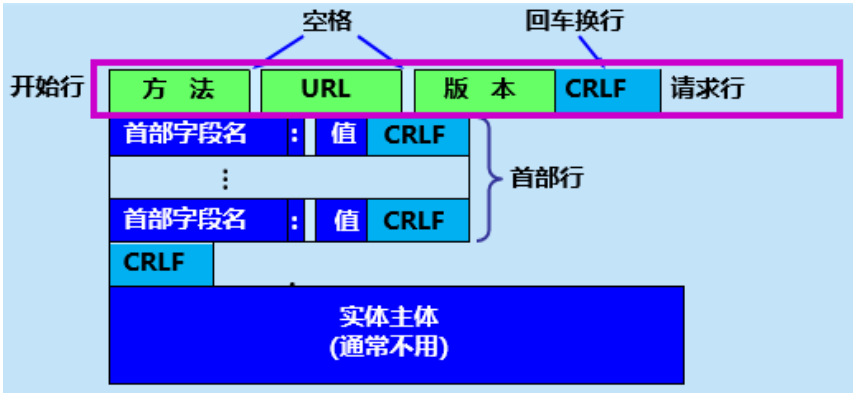
(4) 代理服务器减少了源点服务器的通信量，减小了互联网的时延

3. HTTP的报文结构

(1) 由于HTTP是text-oriented面向正文的，所以每个字段都是ASCII码串，没有固定长度，大致分为开始，首部，主体三部分

(2) 请求报文：客户向服务器发的报文

1)



2) URL是请求资源的URL

3) 版本是HTTP的版本号

4) 首部是指浏览器、服务器、报文主体等信息

5) entity主体在请求报文一般用不到

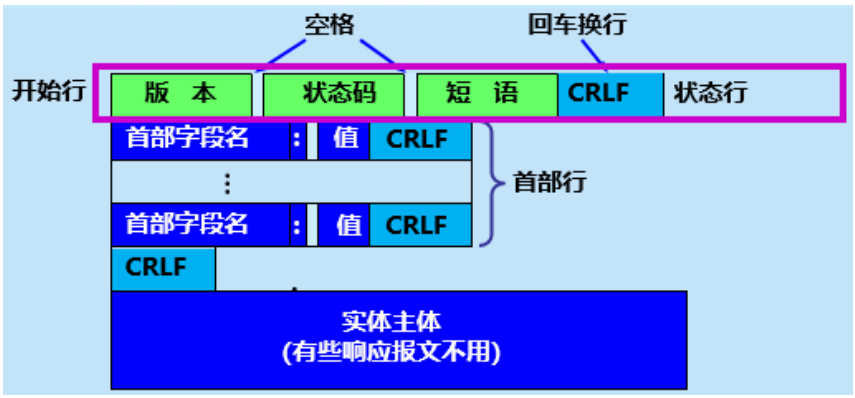
(3) 请求报文中的方法

1) OPTION 请求一些选项的信息

2) GET 请求读取由 URL所标志的信息
3) HEAD 请求读取由 URL所标志的信息的首部
4) POST 给服务器添加信息（例如，注释）
5) PUT 在指明的 URL下存储一个文档
6) DELETE 删除指明的 URL所标志的资源
7) TRACE 用来进行环回测试的请求报文
8) CONNECT 用于代理服务器

(4) 响应报文

1)



2) 短语是解释状态码用的

(5) 响应报文的status code状态码

1) 1xx 表示通知信息的，如请求收到了或正在进行处理
2) 2xx 表示成功，如202接受或知道了
3) 3xx 表示重定向，如301表示要完成请求还必须采取进一步的行动
4) 4xx 表示客户的差错，如400错误语法和404not found
5) 5xx 表示服务器的差错，如服务器失效无法完成请求

(6) 对GET请求，浏览器会一次发送header和data，服务器返回200

(7) 对POST请求，浏览器先发header，待服务器返回100，再发送data

4. 在服务器上存放用户的信息

(1) Cookie：HTTP服务器和客户间传递的状态信息

(2) 当用户访问有Cookie的网站时，服务器会给用户产生一个识别码，并作为索引加入数据库，并在响应报文首部里添加Set-Cookie：号码；浏览器收到该报文后也将该Cookie码放在以后的报文请求的首部中

(3) 之后该用户浏览器再访问该网络时，服务器就能识别出该用户，避免重新输入信息的过程

(4) 虽然Cookie号本身并不会影响用户计算机的信息安全，但也许用户隐私会被网站知道，可以在浏览器Internet选项设置里更改

4- 万维网的文档

1. 超文本标记语言HTML

(1) 并不是应用层协议，只是万维网浏览器使用的语言，RFC2854，由WWW Consortium负责制订

(2) 可以用.txt后缀来更改html文档，但只有改为.html或.htm后才能被浏览器显示成想要的效果

(3) 定义了许多用于排版的tag标签，如<I>表示之后的是斜体，</I>表示之

后的不是斜体，这些在浏览器中会转换成对应的效果

(4) 链接

1) 起点可以是文字，图片

2) 远程链接终点在其他网站的页面

3) 本地链接终点指向本计算机中的某个文件

(5) eXtensible Markup Language可扩展标记语言，不同于HTML目的是显示数据，XML的设计宗旨是传输数据，简单，平台无关，是对HTML的补充，用户界面与结构化的数据分隔开

(6) eXtensible HTML可扩展超文本标记语言，是作为XML的应用而被重新定义的HTML，是WWWC在00年制定的标准

(7) Cascading Style Sheets层叠样式表CSS用于为HTML文档定义布局，HTML用于结构化内容，CSS用于结构化结构化的内容，主要指字体、颜色、边距、高度、宽度、背景等

2. 动态万维网文档

(1) 之前讲的都是static document静态文档，内容在创作完后不会变

(2) dynamic document动态文档指被访问时由服务器端应用程序动态创建的，而在客户端看到的还是静态文档格式的

(3) 服务器端需要增加：一个应用程序，处理浏览器发来的数据，并创建出建立动态文档需要的数据；一个机制，将浏览器发来的数据传给该应用程序，并解释出其返回的数据，建立成HTML文档

(4) Common Gateway Interface通用网关接口CGI，是一种标准，定义了动态文档如何创建，输入数据如何送给应用程序，输出结果如何使用

(5) 符合CGI标准的CGI程序又称为CGI script脚本，因为常备放在/cgi-bin目录下，又被称为cgi-bin脚本

3. 活动万维网文档

(1) server push服务器推送：将所有工作交给服务器，不断更新动态文档

(2) 过多服务器推送程序会造成过多服务器开销，于是又出现了浏览器端负责更新的active document活动文档，由浏览器负责运行服务器返回的活动文档程序副本，生成静态文档

(3) 类似Java的applet小应用程序也可用于描述活动文档程序

5- 万维网的信息检索系统

1. 全文检索搜索与分类目录搜索

(1) search engine搜索引擎主要有全文检索和分类目录两种

1) 全文检索是先收集各网站的信息建立数据库，再从用户输入的关键字，去查询，需要及时更新，如谷歌

2) 分类目录是将网站描述经人工审核后添加到某关键字对应目录的数据库中，如雅虎，新浪等

(2) vertical垂直搜索引擎：针对某一行业知识的上下文，为满足特定人群，返回信息、消息、条目，如购物、旅游、求职、交友

(3) meta元搜索引擎：将用户检索请求发到多个独立搜索引擎中，再整合结

果，速度和智能化、个性化都很强，查全率和查准率也很高

2. Google搜索技术的特点
   (1) 利用率互联网上的相链接的计算机来快速查找每个搜索的答案，核心技术是PageRank网页排名
   (2) 将被更多网站指向的网页视作更重要的，初值假设所有网站重要性相同，用二维稀疏矩阵乘法计算重要性排名，多次迭代，总能收敛到真正的“重要度”

6- weblog博客和microblog微博

7- Social Networking Site社交网站SNS

1)
2)
3)
4)
5)
6)
7) ----------我是底线----------

6电子邮件

2019年6月19日 22:55

-
- 电子邮件

1- 电子邮件概述

1. e-mail电子邮件：用电子设备交换的邮件
   (1) User Agent用户代理：用户与邮件系统的接口，又称电子邮件客户端
   (2) 邮件服务器
   (3) 邮件发送协议和邮件读取协议
2. 重要标准
   (1) Simple Mail **Transfer** Protocol简单邮件传送协议SMTP，RFC5321，只传可打印ASCII码
   (2) 互联网文本报文格式，RFC5322
   (3) Multipurpose Internet Mail Extensions通用互联网邮件扩充
   (4) Post Office Protocol邮局协议和Internet Message Access Protocol网际报文存取协议
3. 用户代理UA的功能
   (1) 撰写：提供编辑信件的环境
   (2) 显示来信内容
   (3) 处理：发送和接收邮件
   (4) 通信：用邮件发送和读取协议进行本地到邮件服务器间的发送或接收
4. 通信过程
   (1) 在UA编辑完邮件并点击发送按钮
   (2) UA成为客户，建立TCP连接，用发送协议，将邮件发给邮件服务器
   (3) 邮件服务器将邮件存进缓存队列，等待发到接收方的邮件服务器，这个等待时间一般远大于路由器转发排队的时间
   (4) 发送方的邮件服务器和接收方的邮件服务器建立TCP连接，并发送邮件（若一直建不了连接，会通知用户代理，因此邮件服务器应尽量保持不间断的运行）
   (5) 接收方的邮件服务器中的发送协议进程收到邮件后存入该收件人的邮箱
   (6) 收件人点开UA，去其邮件服务器收信
5. TCP/IP的电子邮件格式
   (1) 电子邮件包括envelope信封和content内容
   (2) 信封上需要标记收件人e-mail address地址
   (3) 地址格式为：用户名@邮件服务器域名

2- 简单邮件传送协议SMTP

1. 连接建立
   (1) 用熟知端口号25与服务器建TCP连接，服务器返回220service ready
   (2) 客户发送HELLO命令，服务器能接收邮件时回答250OK；否则

421Service not available

(3) SMTP不能使用中间的邮件服务器，连接失败只能等一段时间尝试重连

2. 邮件传送

(1) MAIL FROM：命令及之后的发件人地址开始

(2) 若SMTP服务器返回250OK，说明服务器准备好接受邮件

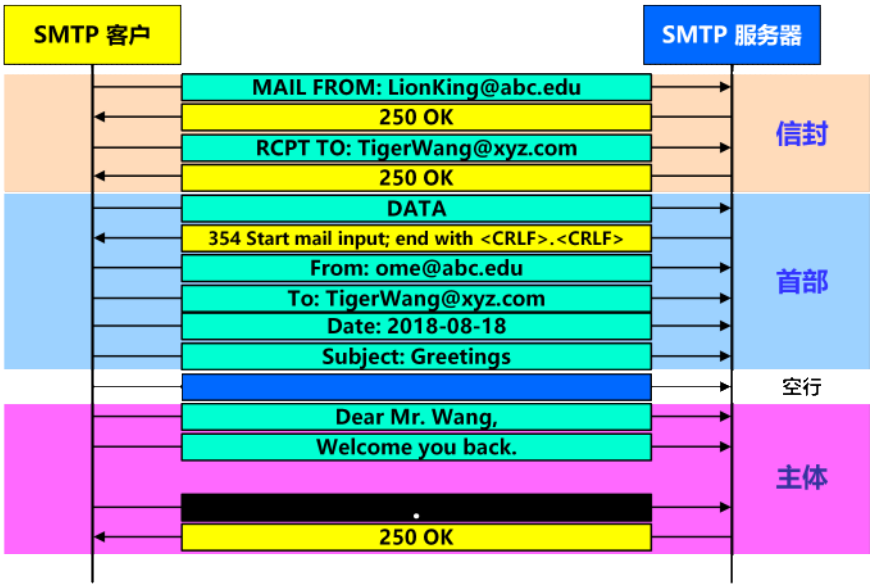
(3) 之后是RCPT TO命令，弄清接收方系统有无该用户，避免浪费通信资源

(4) 若SMTP服务器又返回250OK，说明接收方系统准备好了

(5) 之后是DATA命令表示传送内容，用回车换行.回车换行结束

(6) 这次返回250OK说明发送成功了（但收件人不一定去看了）

(7)



3. 连接释放

(1) 发送完毕后，客户应发送QUIT命令，若服务器返回221服务关闭表示同意释放TCP连接

(2) 以上内容一般都是UA后台完成的，用户看不见

4. eXtended SMTP扩充，RFC5321，08年10月

(1) 客户端鉴别功能、二进制报文、分块传送、国际化地址

(2) 使用了安全传输TLS

(3) 发送的试探报文内容改为EHLO，根据服务器是否回复250判断服务器是否接受XSMTP

3- 电子邮件的信息格式

1. 信封主要是MAIL FROM和RCPT TO

2. RFC3522规定了内容的header首部格式

(1) To:后面填入一个或多个收件人的电子邮件地址

(2) Subject:是邮件的主题。反映了邮件的主要内容，便于用户查找邮件

(3) Cc:表示应给某某人抄送一个邮件副本（Carbon copy）

(4) Bcc暗送：Blind盲腹泻，使收件人不知道自己收到的是副本

(5) From和Date表示发信人的电子邮件地址和发信日期

(6) Reply-To是对方回信所用的地址

3. 内容的主体部分是用户自己撰写的

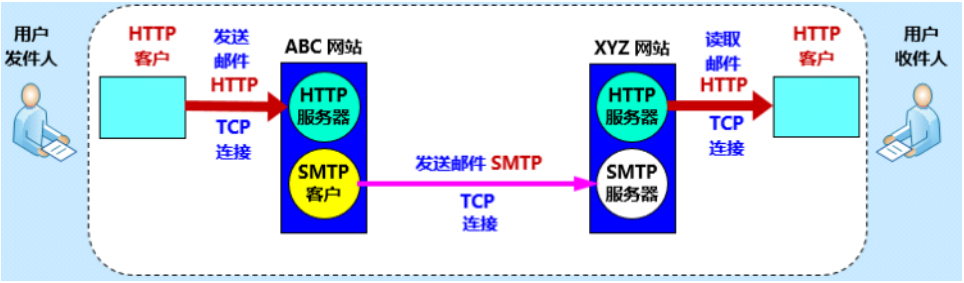
4- 邮件读取协议POP3和网际报文存取协议IMAP4

1. 都使用客户-服务器方式，都基于TCP的实现客户与服务器的通信
2. POP3支持用户鉴别；用户读取整个邮件，服务器随后立即删除该邮件
3. IMAP是联机协议，必须在联网后再读取信件；连接后只下载首部；用户可在IMAP服务器上创建和管理文件夹；可搜索邮件内容；可在不同计算机处理邮件；允许只读一部分

操作位置	操作内容	IMAP	POP3
收件箱	阅读、标记、移动、删除邮件等	客户端与邮箱更新同步	仅在客户端内
发件箱	保存到已发送	客户端与邮箱更新同步	仅在客户端内
创建文件夹	新建自定义的文件夹	客户端与邮箱更新同步	仅在客户端内
草稿	保存草稿	客户端与邮箱更新同步	仅在客户端内
垃圾文件夹	接收并移入垃圾文件夹的邮件	支持	不支持
广告邮件	接收并移入广告邮件夹的邮	支持	不支持

4. SMTP 协议用于发信人的UA向源邮件服务器发送邮件，以及源邮件服务器向目的邮件服务器发送邮件；而 POP 协议或 IMAP 协议则是用户从目的邮件服务器上读取邮件所使用的协议

5- 基于万维网的电子邮件

1. UA软件的缺点：用户计算机内不安装软件就不能使用邮件功能更
2. 万维网电子邮件的优点：不需要安装专用软件，联网就能收发邮件；界面友好
3. 

   (1) 客户端的浏览器和服务器间使用HTTP协议，但服务器间仍是SMTP

6- Multipurpose Internet Mail Extensions通用互联网邮件扩充MIME

1. SMTP缺点：只能传送7位ASCII码；有长度限制；RFC821没被严格遵守（主要指关于回车换行空格和截断的问题）
2. MIME并没有改动SMTP或取代它，而是沿用RFC822，增加了主体结构，定义了非ASCII码的编码规则
   (1) 5个新的首部字段，提供主体信息
      1) MIME-Version：MIME 的版本。若无此行，则为英文文本
      2) Content-Description：这是可读字符串，此邮件的说明
      3) Content-Id：邮件的唯一标识符
      4) Content-Transfer-Encoding：传送时邮件主体使用的编码方法
      5) Content-Type：邮件内容类型 / 子类型

(2) 对多媒体邮件的表示方法进行了标准化

(3) 新定义了传送编码的转换方法

3. Content-Transfer-Encoding内容传送编码

(1) 7位ASCII码：保持不变

(2) quoted-printable

1) 不改变‘=’以外的可打印ASCII码

2) ‘=’和其他编码的字符，每个字节转换成2个16位数对应的ASCII符，再用‘=’开头，多了约两倍空间

(3) base64

1) 每24位按这个格式转换成4个字符：0~25->A~Z；26~51->a~z；52~61->‘0’~‘9’；62->‘+’；63->‘/’

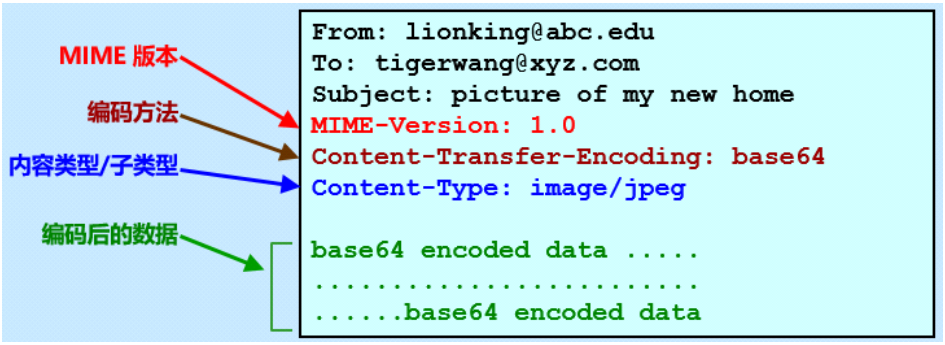
2) 若最后只有一字节或二字节，则用“==”和“=”表示

3) 多了约1/4空间开销

4. 内容类型Content-Type：内容类型type和子类型subtype

内容类型	子类型举例	说明
text（文本）	plain, html, xml, css	不同格式的文本
image（图像）	gif, jpeg, tiff	不同格式的静止图像
audio（音频）	basic, mpeg, mp4	可听见的声音
video（视频）	mpeg, mp4, quicktime	不同格式的影片
model（模型）	vrml	3D模型
application（应用）	octet-stream, pdf, javascript, zip	不同应用程序产生的数据
message（报文）	http, rfc822	封装的报文
multipart（多部分）	mixed, alternative, parallel, digest	多种类型的组合

(1)



1)

2)

3)

4)

5)

6)

7) ----------我是底线----------

6DHCP，SNMP和P2P

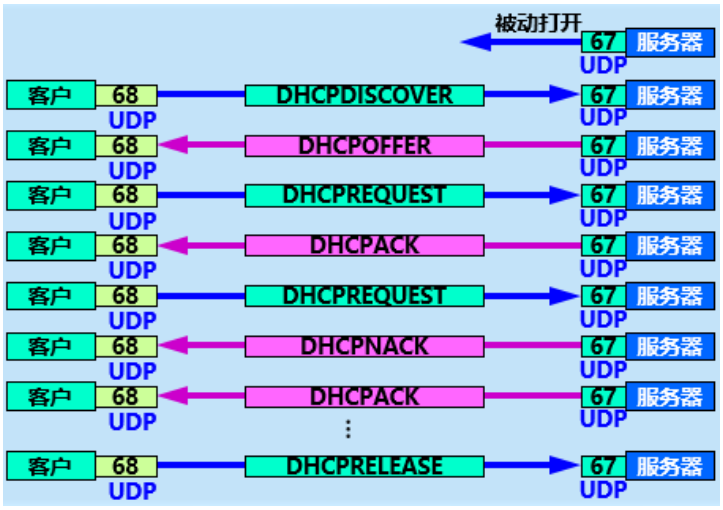
2019年6月20日 14:54

-
- 动态主机配置协议DHCP

1. 协议配置：在协议软件中，给协议参数赋值的动作。是协议软件使用前必做的
   (1) 一般有IP地址、子网掩码、默认路由器IP地址、域名服务器IP地址
   (2) 通常这种信息存储在一个配置文件

1- Dynamic Host Configuration Protocol提供了plug-and-play networking即插即用连网的机制

1. 给运行DHCP服务器软件、位置固定的服务器指派一个永久地址，给运行客户端软件的计算机分配临时地址，RFC2131，RFC2132
   (1) 服务器端口号67
   (2) 客户端端口号68
   (3) 虽然还是草案，但Windows的TCP/IP、属性、自动获得IP地址和自动获得DNS服务器地址选项，其实就是选择使用DHCP
   (4) 据说服务器炸了以后会返回169.255.x.x
2. 基于UDP，使用客户-服务器方式
   (1) 需要IP地址的主机成为客户，自行广播求DHCPDISCOVER（源IP地址暂无，因而全0，目的地址广播，全1）
   (2) 收到发现的服务器返回DHCPOFFER返回回答报文，通知其IP地址
   (3) 客户向服务器发出DHCPREQUEST请求IP
   (4) 本地网络的DHCP服务器在数据库中找该该计算机的配置，并用DHCPACK返回回答报文，不同意分配IP则返回DHCPNACK
   (5) 若找不到，则从IP address pool找找一个空的返回给他
3. relay agent中继代理
   (1) 并不是每个网络都有DHCP服务器，但可以让中继单播代替DHCPDISCOVER找其他网的DHCP服务器
   (2) DHCP服务器再返回DHCPOFFER给中继
4. lease period租用期：DHCP用户分配到IP地址的有限时间
   (1) 一般由DHCP服务器决定，客户也可在DHCPREQUEST报文里提出请求
   (2) 一般客户要记一个计时器，租用期过一半后重新申请
   (3) 或DHCPRELEASE提前结束租用

5. 

-
- 简单网络管理协议SNMP

1- Network Management网络管理的基本概念

1. 网络管理/网管：对硬件、软件和人力的使用、综合与协调，以便对网络资源进行监视、测试、配置、分析、评价和控制，这样就能以合理的价格满足网络的一些需求，如实时运行性能，服务质量等
   (1) 不是指对网络进行行政上的管理
2. 五大功能
   (1) 故障管理：故障检测、隔离和纠正
   (2) 配置管理：初始化网络、并配置网络
   (3) 计费管理：记录网络资源的使用
   (4) 性能管理：估价系统资源的运行状况及通信效率等
   (5) 网络安全管理：对授权机制、访问控制、加密和加密关键字的管理
3. 主要构成
   (1) Network Operations Center网络运行中心/管理站
   (2) 管理程序及管理进程
   (3) manager管理者：管理站硬件或管理程序软件
   (4) administrator管理员：负责管理的人
4. Managed Object被管对象：被管理的硬件及软件
   (1) 每个被管设备（主机、路由器、打印机、网桥等）内都有许多被管对象
   (2) 硬件被管对象主要有网络接口卡等
   (3) 软件被管对象如协议参数的集合
   (4) 被管对象形成一棵对象命名树
5. 网络管理代理程序/agent代理：每个被管设备中负责和管理站通信的程序
   (1) 在管理程序发来的命令和控制下对被管设备采取本地行动
   (2) 网管基本原理：管理某个对象肯定要添加软件或硬件，尽量减少这种添加给对象带来的影响，出于这种考虑，产生了基于UDP的SNMP
   (3) SNMP服务器端口161；客户端端口162
6. SNMP基本功能：监视网络性能、检测分析网络差错、配置网络设备
   (1) 网络正常工作时，SNMP可统计、配置、测试

(2) 网络故障时，SNMP可差错检测、恢复
(3) SNMPv3的安全特性被改进，RFC 3411~3418
(4) proxy agent委托代理：通过协议转换和过滤操作，实现对非SNMP的网络元素进行管理

7. SNMP组成：SNMP本身、Structure of Management Information管理信息结构SMI、Management Information Base管理信息库MIB
   (1) SNMP定义了管理站和代理间交换的分组格式，需要包含管理对象（变量）和状态（值），SNMP负责读取和改变值
   (2) SMI定义命名对象和定义对象类型（范围和长度）的通用规则，和把对象和值编码的规则，确保了语法语义无二义性
   (3) MIB给被管实体创建了命名对象，规定了类型
   (4) SMI建立规则，MIB对变量进行说明，SNMP完成网管动作
      -
      - 应用进程跨越网络的通信

1- 系统调用和应用编程接口

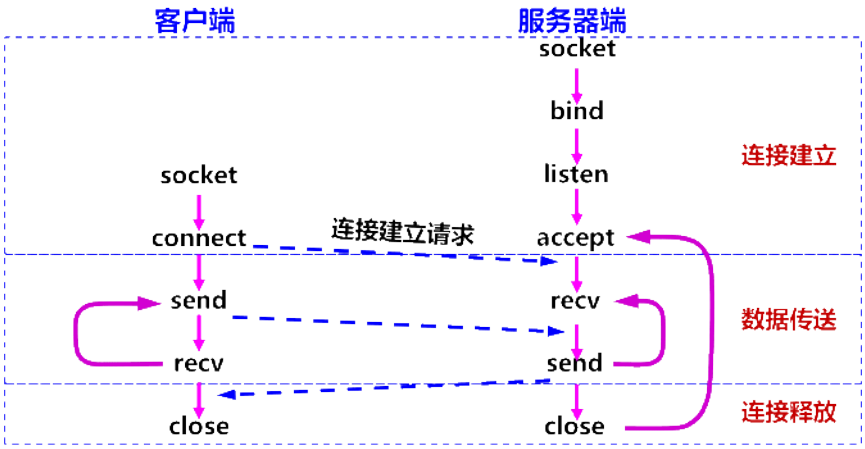
1. 大多数操作系统使用系统调用 (system call) 的机制在应用程序和操作系统之间传递控制权
   (1) 系统调用和一般程序设计中的函数调用非常相似
   (2) 只是系统调用是将控制权传递给了操作系统
2. 系统调用接口实际上就是应用进程的控制权和操作系统的控制权进行转换的一个接口。使用系统调用之前要编写一些程序，特别是需要设置系统调用中的许多参数，因此这种系统调用接口又称为Application Programming Interface 应用编程接口 API
   (1) Berkeley UNIX 操作系统定义了一种 API，它又称为套接字接口 (socket interface)。
   (2) 微软公司在其操作系统中采用了套接字接口 API，形成了一个稍有不同的 API，并称之为 Windows Socket。
   (3) AT&T 为其 UNIX 系统 V 定义了一种 API，简写为 TLI (Transport Layer Interface)。

2- 常用的系统调用

(1) 应用进程需要使用网络进行通信时，就发出系统调用。
(2) 使用 TCP/IP 应用编程接口 API，就可以编写基于互联网的网络应用程序了。
(3) 调用 API 时，用户可以使用 TCP 服务，也可以使用 UDP 等其他服务。

1. TCP 提供面向连接的服务。
   (1) 使用TCP服务需要经历 3 个阶段：连接建立阶段、数据传送阶段、连接释放阶段

(2)



2. 建立连接阶段：
   (1) 当套接字被创建后，它的端口号和 IP 地址都是空的，因此应用进程要调用 bind（绑定）来指明套接字的本地地址。在服务器端调用 bind 时就是把熟知端口号和本地 IP 地址填写到已创建的套接字中。这就叫做把本地地址绑定到套接字。
   (2) 服务器在调用 bind 后，还必须调用 listen（收听）把套接字设置为被动方式，以便随时接受客户的服务请求。（UDP 服务器由于只提供无连接服务，不使用 listen 系统调用。）
   (3) 服务器紧接着就调用 accept（接受），以便把远地客户进程发来的连接请求提取出来。系统调用 accept 的一个变量就是要指明从哪一个套接字发起的连接。 调用 accept 要完成的动作较多。这是因为一个服务器必须能够同时处理多个连接。这样的服务器常称为并发方式 (concurrent) 工作的服务器。
3. 传送阶段：
   (1) 客户和服务器都在 TCP 连接上使用 send 系统调用传送数据，使用 recv 系统调用接收数据。
   (2) 通常客户使用 send 发送请求，而服务器使用 send 发送回答。
   (3) 服务器使用 recv 接收客户用 send 调用发送的请求。客户在发完请求后用 recv 接收回答。
4. 连接释放阶段：
   (1) 一旦客户或服务器结束使用套接字，就把套接字撤消。这时就调用 close 释放连接和撤销套接字。

- ◆
- ◆ P2P应用

1- P2P工作方式概述

(1) P2P 工作方式受到广大网民的欢迎。
(2) 在 P2P 工作方式下，所有的音频/视频文件都是在普通的互联网用户之间传输。
(3) 这种工作方式解决了集中式媒体服务器可能出现的瓶颈问题。
(4) 在互联网流量中，P2P 工作方式下的文件分发已占据了最大的份额，比万维网应用所占的比例大得多。

1. 具有集中目录服务器的 P2P 工作方式
   (1) Napster 最早使用 P2P 技术，提供免费下载 MP3 音乐。
   (2) Napster 将所有音乐文件的索引信息都集中存放在 Napster 目录服务器中。
   (3) 使用者只要查找目录服务器，就可知道应从何处下载所要的MP3文件。
   (4) 用户要及时向 Napster 的目录服务器报告自己存有的音乐文件。
   (5) Napster 的文件传输是分散的，文件的定位则是集中的。
   (6) 缺点：集中目录服务器可靠性差且可能会成为性能瓶颈
2. 具有全分布式结构的 P2P 文件共享程序
   (1) 如Gnutella采用洪泛法查询代替集中目录服务器
   (2) 如eMule服务器保存用户有关信息、提供共享文件夹、鼓励上传文件最多的用户下载优先级高
   (3) BitTorrent比特洪流BT
      1) 比特洪流：为所有参与文件分发的对等用户建立一个torrent洪流
      2) chunk文件块：下载的文件的数据单元（长度固定）
      3) tracker追踪器：各洪流的基础设施结点，记录洪流中的对等方
      4) neighboring peers相邻对等方：在追踪器帮助下建立TCP连接的对等方
   (4) BT机制
      1) 文件块请求算法：rarest first，即拥有者最少的文件块优先请求
      2) 文件块发送算法：10秒内速率最快的对等方，称其为unchoked已疏通无障碍对等方，优先发送
3. P2P文件分发的分析
   (1) 设服务器上传速度us，用户上传下载速度ui，di，文件大小F，数量N
   (2) 则C/S模式N个用户最少下载时间Tcs=max{NF/us, F/min{di}}
      1) 即瓶颈为服务器接入链路或最慢主机接入链路
   (3) P2P模式Tp2p>max{F/us, F/min{di}, NF/sum{ui}}
      1) 不取等号是因为用户获得文件块后不能立刻成为上传方
      2) N足够大时只需考虑最后一项
4. 在P2P对等方中搜索对象
   (1) Kid：资源名关键字
   (2) Nid：存放资源的结点的IP地址，可能附带端口号
   (3) Distributed Hash Table分布式散列表DHT，如Chord环Pastry、CAN(Content Addressable Network)、Kademilia等
   (4) 一般会用到哈希算法建立Kid、Nid到其散列值的散列表
   (5) 如Chord环形结构配合指针表加速查找
      1)
      2)
      3)
      4)

5)
6)
7)
8)
9)
10) ----------我是底线----------

7网络安全

2019年6月20日 15:18

-
- 网络安全问题概述

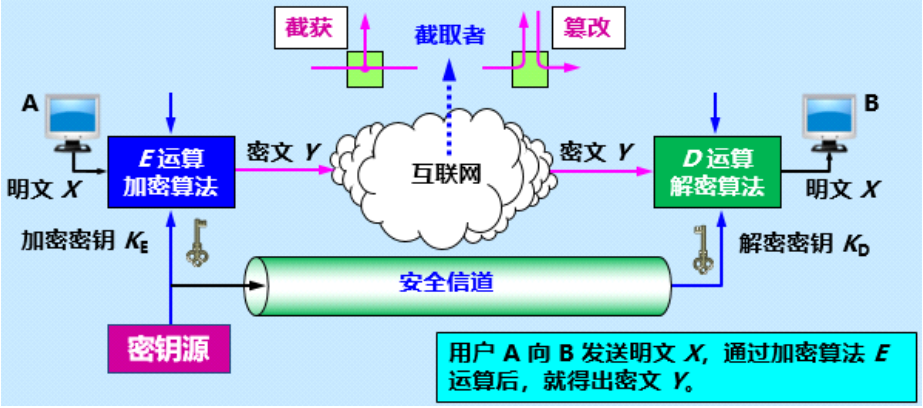
1- 计算机网络面临的安全性威胁

1. 被动攻击/窃听/截获：观察和分析协议数据单元PDU，但不干扰信息
   (1) 一般不易理解内部信息，但可以观察协议控制信息，了解通信实体地址和身份，研究PDU长度和通信频度，因而又被称traffic analysis流量分析
   (2) 战争时期，发现某处大量异常通信量就能猜到指挥所位置
2. 主动攻击
   (1) 篡改：修改报文流
   (2) rogue program恶意程序：
      1) virus病毒：修改其他程序，把自己或变种复制进去“传染”
      2) worm蠕虫：用通信功能将自己发到其他结点并自启动
      3) Trojan horse特洛伊木马：藏在其他程序里，再偷偷复制出来
      4) logic bomb逻辑炸弹：定时删去系统文件
      5) backdoor knocking后门入侵：用系统漏洞，通过网络入侵系统
      6) 流氓软件：强制安装，难卸载，劫持浏览器，弹广告，收集用户信息，恶意卸载，恶意捆绑等
   (3) Denial of Service拒绝服务DoS/网络带宽攻击/连通性攻击：向服务器发送大量分组，直至瘫痪。若互联网上不同计算机攻击同一网站，又称Distributed分布式拒绝服务
   (4) 交换机中毒poisoning：伪造不存在的MAC地址，使交换表满
3. 被动攻击难检测，但主动攻击可采取措施检测，因此，网络安全的目标：
   (1) 防止被分析出报文内容和流量
   (2) 防止恶意程序
   (3) 检测篡改和拒绝服务
4. 实现方法：加密解决被动攻击，加密和鉴别解决主动攻击

2- 安全的计算机网络

1. 保密性：只有发送接收方看得懂
2. 端点鉴别：能确定发送接收方的真实身份
3. 信息完整：不被篡改
4. 运行安全：access control访问控制权限，尤其在multilevel security多级安全下更重要

3- 数据加密模型

1. 

(1) 加密encryption；解密deciphering

(2) 如果不论截取者获得了多少密文，但在密文中都没有足够的信息来唯一地确定出对应的明文，则这一密码体制称为无条件安全的，或称为理论上是不可破的

(3) 如果密码体制中的密码不能被可使用的计算资源破译，则这一密码体制称为在计算上是安全的

- **加密和解密用的密钥*K*(key)是一串秘密的字符串（即比特串）。**
- **明文通过加密算法 *E* 和加密密钥 *K* 变成密文：**

$$Y = E_K(X) \qquad (7\text{-}1)$$

2.
- **接收端利用解密算法 *D* 运算和解密密钥 *K* 解出明文 *X*。解密算法是加密算法的逆运算。**

$$D_K(Y) = D_K(E_K(X)) = X \qquad (7\text{-}2)$$

- **加密密钥和解密密钥可以一样，也可以不一样。**
- **密钥通常由密钥中心提供。**
- **当密钥需要向远地传送时，一定要通过另一个安全信道。**

(1) cryptography密码编码学是密码体制的设计学。

(2) cryptanalysis密码分析学则是在未知密钥的情况下从密文推演出明文或密钥的技术

(3) cryptology密码编码学与密码分析学合起来即为密码学

-
- 两类密码体制

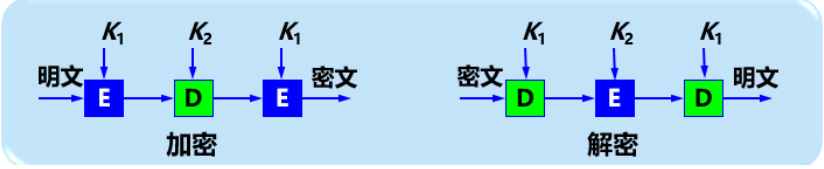
1- 常规/对称密钥密码体制：加密密钥与解密密钥是相同的密码体制

1. Data Encryption Standard数据加密标准DES：保密性仅取决于密钥的保密，算法是公开的

(1) 每64位一组，处理加密，实际密钥长度56位，8位用于奇偶校验

(2) triple DES：用一个密钥加密，再用另一个密钥解密，然后再使用第一个密钥加密，即

$$Y = DES_{K1}(DES^{-1}_{K2}(DES_{K1}(X))) \qquad (7\text{-}3)$$

1) 

(3) 广泛用于网络、金融、信用卡等

(4) NIST最终选择比利时的Rijndael算法为Advanced Encryption

StandardAES，以取代DES

2- 公钥密码体制：使用不同的加密密钥与解密密钥

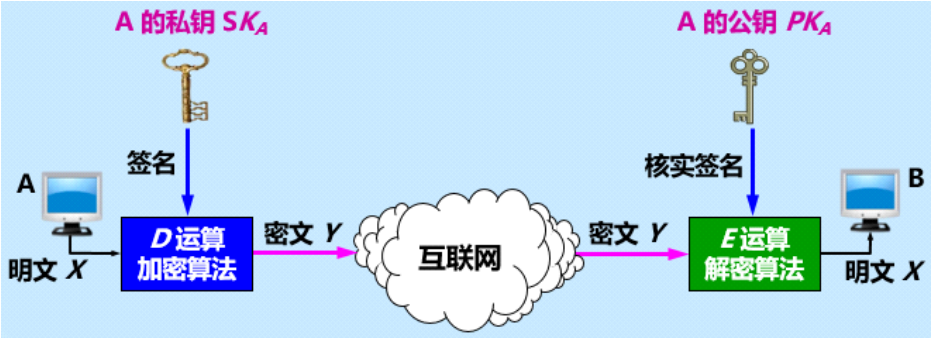
1. 是“由已知加密密钥推导出解密密钥在计算上是不可行的”密码体制
2. 加密算法E和解密算法D都是公开的，加密密钥也是public key公钥，而解密密钥是secret key私钥/秘钥，是唯一需要保密的（注意，根据公钥猜不到秘钥）
   (1) 用公钥加密，私钥解密，即实现了多对一保密
   (2) 用私钥加密，公钥解密，及实现了数字签名
   (3) 避免大量密钥分配的问题，和实现对数字签名的要求是公钥体制的出现原因
3. 因为加密方法安全性仅取决于密钥长度及攻破所需计算量
   (1) 公钥体制算法开销较大，密钥分配协议也很麻烦
4. 加密方法：发送者 A 用 B 的公钥 PKB 对明文 X 加密（E 运算）后，接收者 B 用自己的私钥 SKB 解密（D 运算），即可恢复出明文：
   (1) $D_{SK_B}(Y)=D_{SK_B}(E_{PK_B}(X))=X$
   (2) 注意加密密钥不能用来解密
   (3) $D_{PK_B}(E_{PK_B}(X)) \neq X$
   (4) 但加密和解密作为互逆运算，可以调换顺序
   (5) $E_{PK_B}(D_{SK_B}(X))=D_{SK_B}(E_{PK_B}(X))=X$

-
- 数字签名

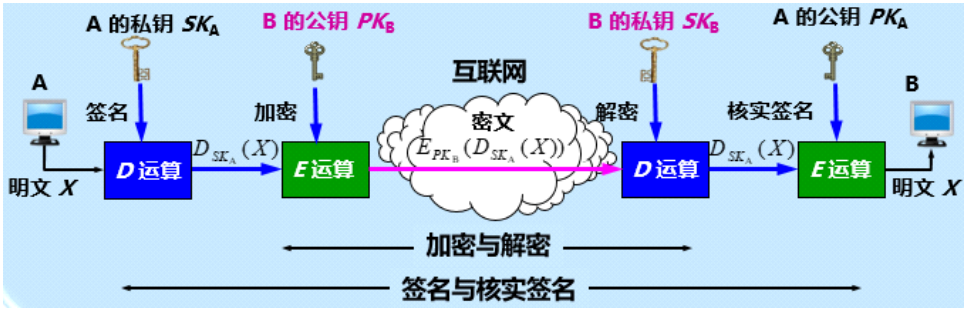
1- 数字签名目的

1. 报文鉴别——接收者能够核实发送者对报文的签名（证明来源）；
2. 报文的完整性——发送者事后不能抵赖对报文的签名（防否认）；
3. 不可否认——接收者不能伪造对报文的签名（防伪造）

2- 基于公钥的简单数字签名

1. 鉴别来源：除 A 外没有别人能具有 A 的私钥，所以没有别人能产生这个密文。因此 B 相信报文 X 是 A 签名发送的
2. 内容完整：若 A 要抵赖曾发送报文给 B，B 可将明文和对应的密文出示给第三者。第三者很容易用 A 的公钥去证实 A 确实发送 X 给 B。
3. 不可否认：若 B 将 X 伪造成 X'，则 B 不能在第三者前出示对应的密文。这样就证明了 B 伪造了报文
4. 

3- 具有保密性的基于公钥的数字签名

1. 用接收方的公钥做第二次加密
2. 

-
- 鉴别

1. Authentication鉴别：
   (1) 实体鉴别：验证通信对象非冒充者
   (2) 报文鉴别：确认报文完整，包括确认发送者无误
2. 与加密是不同的概念
3. 与authorization也是不同的概念，授权目的是确认该过程是否被允许

1- 报文鉴别

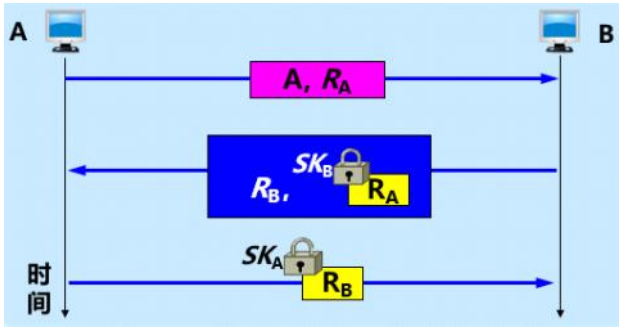
(1) 数字签名给计算机很大负担，不常用于鉴别

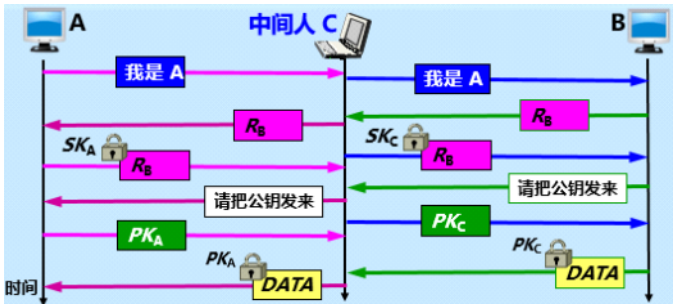
1. cryptographic hash function密码散列函数
   (1) 散列函数：对任意长的输入内容，输出一个长度固定的值
   (2) 散列函数是多对一的one-way单向函数，从散列函数值是可以伪造出伪原文的
   (3) 将固定位数的散列函数值拼接在报文后，方便检验，之前的checksum检验和就是这个思想的应用
2. Message Digest5报文摘要，RFC1321，1991年
   (1) 把二进制报文长度模2^64，求得64位余数，添加在报文后
   (2) 在报文后和长度余数间添加1个1和不超过511个0，使新长度为512倍数
   (3) 每512位，给报文分组一次，再给每组分为4个128位，再分成4个32位
   (4) 再给每小组算不同的散列函数值
   (5) 经王小云证明，一小时内能伪造出原文
3. Secure Hash Algorithm安全散列算法，稍慢但更安全
   (1) 按512位分组，算报文摘要值
   (2) 将分组和报文摘要值结合，产生报文摘要的下一个中间结果，直至完毕
   (3) 扫描5遍，使抗穷举性更高
4. Message Authentication Code报文鉴别码MAC
   (1) 只拼接一个散列函数值并不能仿伪造原文，最好对散列值再做一次加密

2- 实体鉴别

(1) 报文鉴别对每个报文都要鉴别发送者
(2) 实体鉴别只需在通信前验证一次

1. 最简单的实体鉴别：用只有双方知道的密钥做对称加密
   (1) replay attack重放攻击：截获A的报文，伪装成是C发出该报文的假象
   (2) IP欺骗：C假装自己的IP地址和A一样，让B彻底看不出C不是A

2. nonce不重数：不重复使用的大随机数，即一次一数
   (1) 通过每次重新对不重数用私钥加密，再塞入报文，应对重放攻击（中间人没有私钥，不能加密，相当于不能签名，实现鉴别）
   (2) 

3. man-in-the-middle attack中间人攻击
   (1) 

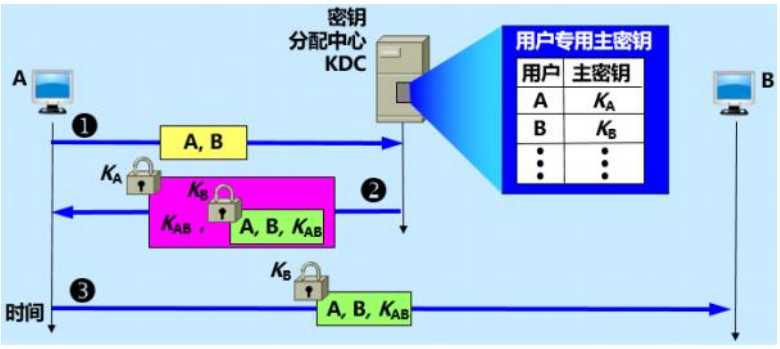
   (2) A 向 B 发送“我是 A”的报文，并给出了自己的身份。此报文被 “中间人” C 截获，C 把此报文原封不动地转发给 B。B 选择一个不重数 RB 发送给 A，但同样被 C 截获后也照样转发给 A
   (3) 中间人 C 用自己的私钥 SKC 对 RB 加密后发回给 B，使 B 误以为是 A 发来的。A 收到 RB 后也用自己的私钥 SKA 对 RB 加密后发回给 B，中途被 C 截获并丢弃。B 向 A 索取其公钥，此报文被 C 截获后转发给 A
   (4) C 把自己的公钥 PKC 冒充是 A 的发送给 B，而 C 也截获到 A 发送给 B 的公钥 PKA
   (5) B 用收到的公钥 PKC （以为是 A 的）对数据加密发送给 A。C 截获后用自己的私钥 SKC 解密，复制一份留下，再用 A 的公钥 PKA 对数据加密后发送给 A
   (6) A 收到数据后，用自己的私钥 SKA 解密，以为和 B 进行了保密通信。其实，B 发送给 A 的加密数据已被中间人 C 截获并解密了一份。但 A 和 B 却都不知道
      -
      - 密钥分配

1. 密钥管理：密钥的产生、分配、注入、验证、使用
   (1) 密钥分配/分发：网外分配、网内分配

1- 对称密钥的分配

1. 对称密钥网内分配难题
   (1) n个人互相通信，需要O（n方）的密钥数量
   (2) 互联网内分配密钥肯定需要加密，这个加密的密钥又如何分配
2. Key Distribution Center密钥分配中心KDC
   (1) 是一种大家都信任的机构，负责给想秘密通信的用户临时分配一个密钥

3. 用户A和B都是KDC事先登记的用户，已在KDC服务器事先分配到了与KDC通信时的master key主密钥KA和KB

4. 

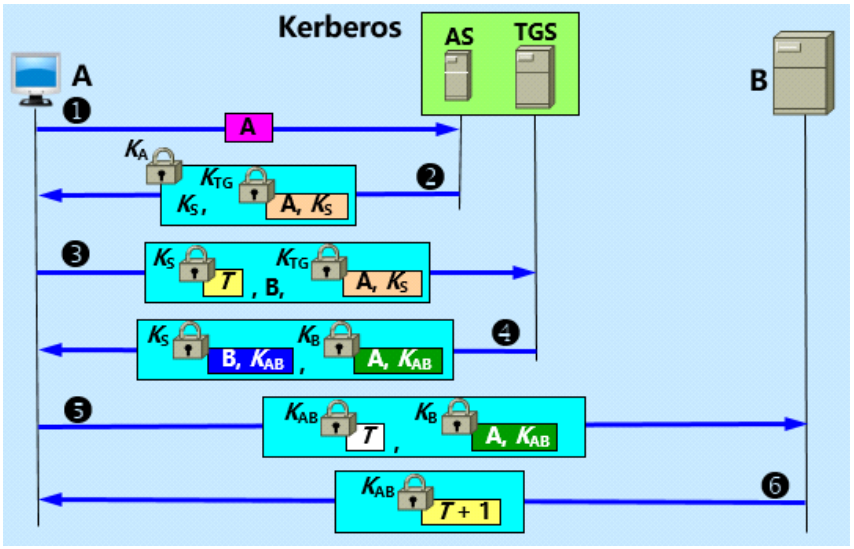
5. 为防止被重放攻击，还可在报文内加入时间戳

6. 主密钥KA和KB应定期更换，减少被破译的机会

1. Kerberos V5协议

    1) Kerbero同时是个KDC

    2) 使用AES加密

    3) 

    (1) A 用明文（包括登记的身份）向鉴别服务器 AS 表明自己的身份

    (2) AS 向 A 发送用 A 的对称密钥 KA 加密的报文，这个报文包含 A 和 TGS 通信的会话密钥 KS，以及 AS 要发送给 TGS 的票据（这个票据是用 TGS 的对称密钥 KTG 加密的）

    (3) A 向 TGS 发送三个项目：

        1) 转发鉴别服务器 AS 发来的票据

        2) 服务器 B 的名字。这表明 A 请求 B 的服务。请注意，现在 A 向 TGS 证明自己的身份并非通过键入口令（因为入侵者能够从网上截获明文口令），而是通过转发 AS 发出的票据（只有 A 才能提取出）。票据是加密的，入侵者伪造不了

        3) 用 KS 加密的时间戳 T 。它用来防止入侵者的重放攻击

    (4) TGS 发送两个票据，每一个都包含 A 和 B 通信的会话密钥 KAB 。给 A 的票据用 KS 加密；给 B 的票据用 B 的密钥 KB 加密。请注意，现在入侵者不能提取 KAB，因为不知道 KA 和 KB 。入侵者也不能重放步骤（3），因为入侵者不能把时间戳更换为一个新的（因为不知道 KS ）

    (5) A 向 B 转发 TGS 发来的票据，同时发送用 KAB 加密的时间戳 T

1) 为防止被重放，Kerberos要求所有主机在时钟上松散的同步，即误差不要超过5分钟

(6) B 用时间戳 T+1 来证实收到了票据。B 向 A 发送的报文用密钥KAB 加密

1) 以后，A 和 B 就使用 TGS 给出的会话密钥 KAB 进行通信

1- 公钥的分配

1. Certification Authority认证中心CA：将公钥与其对应实体进行binding绑定
   (1) 一般是政府出资建立的机构
   (2) 任何用户都能从可信的地方获得CA的公钥，验证某个公钥是否被某个实体拥有
2. 每个实体都有CA发来的certificate证书，里面有公钥及用户标识
   (1) 此证书被CA数字签名了，不可伪造
   (2) ITU-T制定了X.509协议标准，描述整数结构
   (3) IETF对X.509做出了少量改动，给出了Public Key Infrastructure公钥及出结构PKI

版本号	区分 X.509 不同版本
序列号	CA 发放，唯一
签名算法	签署证书所使用的算法和参数
发行者	CA 的 X.509 名字
有效期	包括起始时间和终止时间
主体名	证书持有者的名称及有关信息
公钥	有效的公钥及其使用方法
发行者 ID	任选，唯一，标识发行者
主体 ID	任选，唯一，标识证书持有者
扩展域	扩充信息
认证机构签名	用 CA 私钥对证书签名

3. CA证书过期或吊销
   (1) 用户私钥被泄漏
   (2) 用户不再被CA认证
   (3) CA前述用户证书的私钥被泄漏
   (4) 以上三种情况会使CA证书被吊销，CA需建立并维护一个证书吊销列表
      1)
      2)
      3)
      4)
      5) ----------我是底线----------